# R

## R&D (Research & Development) Intensity

► R&D Investment as a Percentage of GDP

## R&D Investment as a Percentage of GDP

Panagiotis Liargovas
Department of Economics, University of Peloponnese, Tripolis, Greece

### Synonyms

R&D (Research & Development) intensity

### Definition

This is a measure of a country's attitude towards innovation activities. According to Frascati Manual (2002), R&D is considered as a "creative work undertaken on a systematic basis in order to increase the stock of knowledge of man, culture and society, and the use of his stock of knowledge to devise new applications."

### Description

According to OECD (Organization for Economic Cooperation and Development) statistics, R&D investment as a percentage of GDP was on average 2.33 % in OECD countries. The intensity of expenditures in R&D varies significantly among OECD countries. Sweden is the country spending the most (3.75 %) followed by Finland (3.73 %), Japan (3.42 %), and Korea (3.37 %). Mexico, Slovak Republic Poland, and Turkey have the lowest R&D intensity (0.38, 0.47, 0.58, 0.61, and 0.73 %, respectively). Finland and Iceland are the countries that have increased the most their R&D intensity between 1995 and 2005 (OECD, 2007). ► Regional disparities within countries are even larger than among countries. The United States, Sweden, and Korea show that largest disparities in R&D investment as a percentage of GDP across regions. Ireland, together with Greece, the Slovak Republic, Belgium, and Portugal, displays minor differences in R&D investment as a percentage of GDP among regions. It appears that countries with high R&D intensity also exhibit higher internal dispersion (OECD, 2009).

### Discussion

It is often assumed that greater investment in R&D will lead to greater applied research and to an increase in the number of inventions. This linear perception of the innovation process places R&D investment as the key factor behind technological progress and, eventually, ► economic growth. The implications of this approach are that the higher the investment in R&D, the higher the innovative capacity and the higher the economic growth

A.C. Michalos (ed.), *Encyclopedia of Quality of Life and Well-Being Research*,
DOI 10.1007/978-94-007-0753-5, 

(Blaug, 1978). The history of growth suggests that technological change is a major determinant of growth. In his seminal article on the growth of the US economy, (Solow, 1956) showed that only about 12.5 % of the increase in labor productivity between 1909 and 1949 was due to an increase in capital intensity. The rest was attributable to "technical change." The experiences of the newly industrialized countries (NICs) have often been used to justify the salience of technological change in the growth process.

Growth economists seem to agree that economic growth can only be sustained through improvements in technology. An important channel for technological advancement is purposeful research and development activities.

More Specifically, R&D activities have a positive impact on:

(a) FDI (Foreign Direct Investment)
(b) Exports
(c) Manufacturing

### R&D and FDI

It is generally believed that foreign investors prefer to locate in economies with an abundant supply of scientists, engineers, and highly skilled workforce. In other words, they prefer economies that are capable of absorbing and assimilating complex technologies. Since FDI is strongly correlated with growth, it therefore implies that the surfeit of scientists and engineers in a country would spur growth through the attraction of larger flows of FDI. One of the reasons Ireland was able to attract large inflows of high-tech FDI in 1990s was the relatively large number of engineers and other highly trained people it could provide.

### R&D and Exports

Export performance depends both on capital intensity and technological capability. Strong capital intensity enables a country to produce high-tech or innovative products that can compete effectively on international markets. Technological capability also increases total-factor productivity, which subsequently enables a country to become cost-competitive abroad. The experiences of the NICs of South East Asia have shown that export-led industrialization strategies were effective mechanisms for promoting economic growth in 1980s (Krueger, 1997).

### R&D and Manufacturing Value Added

Capital intensity and technological capability generate significant levels of manufacturing value added, which has the effect of speeding up economic growth. The success of firms to develop skills, capabilities, and technical support has led to significant and effective industrialization many countries during the 1990s.

## Cross-References

▶ Economic Growth
▶ Regional Disparities

## References

Blaug, M. (1978). *"Adam Smith and David Ricardo", economic theory in retrospect*. Cambridge, MA: Cambridge University Press.

Krueger Anne, O. (1997). Trade policy and economic development: How we learn. *The American Economic Review, 87*(1): 1–22.

OECD. (2002). Frascati manual, Paris: OECD, available at: www.oecd.org/sti/frascatimanual.

OECD. (2007). *Science and industry scoreboard*. Paris: OECD.

OECD. (2009). *Regions at glance*. Paris: OECD.

OECD. Regional database, http:/stats.oecd.org/WBOS, theme: Regional statistics.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics (The MIT Press) 70*(1): 65–94.

# Racial Discrimination

▶ Xenophobia
▶ Infant Well-Being, Segregation, and Race

# Racism

▶ Cultural or Ethnic Homogeneity Preference Index
▶ Disadvantaged Populations
▶ General Fair Treatment Index

- ▶ Heterogeneous Socializing Network Index
- ▶ Heterogeneous Social Support Index
- ▶ Index of Inegalitarianism
- ▶ Personal Fair Treatment Index
- ▶ Xenophobia
- ▶ Xenophobia Index: Michalos and Zumbo

## Radical Positivism

- ▶ Comparative Analysis

## Radioactive Waste

- ▶ Nuclear Waste

## Radius-Adjusted Trust

- ▶ Trust

## Random Coefficient Modeling

- ▶ Hierarchical Linear Modeling

## Random Coefficients Regression

- ▶ Mixed Effects Models

## Random Effect Modeling

- ▶ Hierarchical Linear Modeling

## Random Effects Regression for Panel Data

Luigi Benfratello
Dipartimento di Scienze Economiche e Statistiche, Università di Napoli Federico II, Napoli, Italy

### Synonyms

Fixed effects models; GLS; Multilevel models; Regression

### Definition

Random effects regression is suited for longitudinal or panel data. The availability of repeated observations on the same units allows the researcher to enrich the model by inserting an additional term in the regression, capturing individual-specific, time-invariant factors affecting the dependent variable but unobserved to the econometrician. Generalized Least Squares estimators of the parameters of such a model are more efficient than those obtained in the simpler model neglecting these unobserved factors.

### Description

The random effects (RE, henceforth) regression model is used both in econometrics, where this model is often opposed to the fixed effects (FE, henceforth) regression model (see below), and in statistics, where this model is a special case of multilevel models. The common context of application of both literatures is the one of longitudinal (or panel) data, i.e., data where several subjects (households, firms, regions) are observed two or more time periods. In multilevel modeling, panel data can be viewed as having a hierarchical structure with measurements (level 1 units) nested into individuals (level 2 units).

The response variable is supposed to be a function of some observables (whose parameters are usually – although not necessarily – supposed to be constant over time and individuals) and some unobservables. In the simplest case, some of these unobservables are specific to the individual/time observations, and others are constant over time at the individual level. The so-called "one-way" error component model is therefore:

$$y_{it} = \beta_0 + \mathbf{x}_{it}\boldsymbol{\beta_1} + \underbrace{\mu_i + \varepsilon_{it}}_{\eta_{it}} \tag{1}$$

where $y_{it}$ is the (scalar) dependent variable for cross-sectional unit $i$ in period $t$, $\mathbf{x}_{it}$ is a $1 \times K$ vector of independent variables observed for unit $i$ in period $t$, $\beta_0$ is the overall intercept, $\boldsymbol{\beta_1}$ is a $K \times 1$ vector of parameters, and $\eta_{it}$ is an error or disturbance term specific to unit $i$ in period $t$, decomposed into a unit-specific and time-invariant component, $\mu_i$, and an observation-specific error, $\varepsilon_{it}$. The peculiarity of the model is the presence of the term $\mu_i$ which captures unobserved heterogeneity at the individual level, i.e., factors constant over time but unobserved to the econometrician. On the one hand, these factors make two observations for the same individual $i$, $y_{it}$ and $y_{is}$, more similar, conditional on the regressor, than two observations for two different individuals $i$ and $j$, $y_{it}$ and $y_{js}$, a very likely situation in panel data context. On the other hand, it leads–as we will see–to estimators different from standard Ordinary Least Squares (OLS, henceforth). The model also specifies the moments (or the distribution) for both $\mu_i$ and $\varepsilon_{it}$. The simplest case is $\varepsilon_{it} \sim iid\left(0,\ \sigma_\varepsilon^2\right)$, $\mu_i \sim\ iid\left(0, \sigma_\mu^2\right)$, both independent of each other and of $\mathbf{X_i}$, the full $T \times K$ matrix of covariates for individual $i$.

The composite nature of the error term $\eta_{it}$ leads to a particular form of its variance-covariance matrix, which highlights a dependence over time of observations $y_{it}$. The $T \times 1$ vector $\boldsymbol{\eta_i}$ has constant variance but non zero correlation because of the time invariant component $\mu_i$:

$$Cov(\boldsymbol{\eta}_i) = E(\boldsymbol{\eta}_i\boldsymbol{\eta}_i') = \boldsymbol{\Omega} = \begin{bmatrix} \sigma_\mu^2 + \sigma_\varepsilon^2 & \sigma_\mu^2 & \dots & \sigma_\mu^2 \\ \sigma_\mu^2 & \sigma_\mu^2 + \sigma_\varepsilon^2 & \dots & \sigma_\mu^2 \\ \vdots & \vdots & & \vdots \\ \sigma_\mu^2 & \sigma_\mu^2 & \dots & \sigma_\mu^2 + \sigma_\varepsilon^2 \end{bmatrix} \tag{2}$$

In this case, OLS regression on the pooled data – although consistent – is not efficient, as it neglects the particular covariance structure of the composite error $\boldsymbol{\eta_i}$ (not spherical as it contains the constant component $\mu_i$). The RE estimator is the Generalized Least Squares (GLS, henceforth) estimator which takes into account the structure of the error in order to obtain an efficient estimator:

$$\widehat{\beta}_{RE} = \left(\sum_{i=1}^{N} \mathbf{X'}_i\boldsymbol{\Omega}^{-1}\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{N} \mathbf{X'}_i\boldsymbol{\Omega}^{-1}\mathbf{y}_i\right) \tag{3}$$

The transformation induced by $\boldsymbol{\Omega}^{-1}$ "quasi-demeans" the data, as it can be proven that the transformed variables are $y_{it} - \theta\bar{y}_i$ and $\mathbf{x}_{it} - \theta\bar{\mathbf{x}}_i$, where $0 \leq \theta \leq 1$ and is a function of $\sigma_\mu^2$, $\sigma_\varepsilon^2$ and $T$. Therefore, a fraction $\theta$ of the individual mean is subtracted to each observation in order to obtain a spherical covariance matrix in the transformed model by taking into account the time invariant component of $\boldsymbol{\Omega}$. When $\theta = 0$, no transformation applies and the efficient GLS estimator is simple OLS on the original data. When $\theta = 1$, the "within-group" transformation applies and the efficient GLS estimator is the FE estimator.

To implement GLS we need to calculate $\boldsymbol{\Omega}$ which in turn requires estimates of the variances $\sigma_\mu^2$ and $\sigma_\varepsilon^2$. Several procedures have been suggested in the literature (see Baltagi, 2008, Chap. 2 for a discussion). They are asymptotically equivalent although they lead to different estimates in finite samples. If the researcher is willing to impose a specific distribution for $\mu_i$ and $\varepsilon_{it}$, Maximum Likelihood estimator of the parameter vector for model (1) $\left(\beta_0, \boldsymbol{\beta_1}, \sigma_\varepsilon^2, \sigma_\mu^2\right)$ can be

used (see Arellano, 2003, Chap. 2; Baltagi, 2008, Chap. 2; Hsiao, 2003, Chap. 3).

The model can be generalized along several directions. The assumptions of homoskedasticity of both $\mu_i$ and $\varepsilon_{it}$ and the lack of autocorrelation of the latter can be relaxed, leading to GLS estimators which use more general covariance matrices than (3) (see Baltagi, 2008, Chap. 5 and Hsiao, Chap. 3). Furthermore, factors common to all individuals in a given time period can be included, leading to the so-called "two-way" error component model (or "two-way crossed effect" models in the multilevel literature):

$$y_{it} = \beta_0 + \mathbf{x}_{it}\boldsymbol{\beta}_1 + \underbrace{\mu_i + \lambda_t + \varepsilon_{it}}_{\eta_{it}} \qquad (4)$$

whose estimation proceeds along the same lines – apart from technical details – as the "one-way" model (see Baltagi, 2008, Chap. 3).

The RE Regression model has been widely used in economics and in social sciences, due to the availability of panel data and the need to include in the model unobserved heterogeneity. The RE estimator has the great advantage of being more efficient than the FE estimator under the above hypotheses. However, its validity has been questioned as the assumption that the time-invariant component, $\mu_i$, is uncorrelated with the covariate matrix $\mathbf{X_i}$ contrasts with the fact that most of the times the covariates are deliberately chosen by individuals on the basis of some factors unobserved to econometrician. Without this assumption, the RE Estimator is inconsistent. This has led to the wide use of the FE Regression model which does not specify any form of relationship between $\mu_i$ and $\mathbf{X_i}$ simply because the within-group transformation above eliminates $\mu_i$ from the estimating equation. The RE model has also been widely used in social sciences, applications being crime in some geographical areas, school achievements and peer effects, voting behavior, smoking attitude, job satisfaction. In social sciences, the RE model is still very widely used, probably because covariates are often measured at the aggregate (and not individual) level and/or less emphasis is put on the possible correlation between unobserved heterogeneity and included regressors. Furthermore, although FE estimator is the standard solution in case of correlation between time-invariant component $\mu_i$ and the covariate matrix $\mathbf{X_i}$, an alternative is to estimate a RE model with individual means of the covariates as additional regressors (see Baltagi, 2008, pp. 133–134). As shown by Mundlak (1978), the ensuing GLS estimator is equivalent to the FE estimator. This approach has the advantage of estimating the effect of time-invariant covariates and is an additional reason why RE models are widely used in social sciences.

## Cross-References

- ▶ Cross-Classified Hierarchical Linear Modeling
- ▶ Hierarchical Linear Modeling
- ▶ Longitudinal Data Analysis

## References

Arellano, M. (2003). *Panel data econometrics*. Oxford: Oxford University Press.

Baltagi, B. H. (2008). *Econometric analysis of panel data* (4th ed.). New York: Wiley.

Hsiao, C. (2003). *Analysis of panel data* (2nd ed.). Cambridge: Cambridge University Press.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica, 46*, 69–86.

## Random Error

- ▶ Measurement Error

## Random Variable

- ▶ Estimator

## Randomized Clinical Trial

- ▶ Experimental Design
- ▶ Factorial Design

## Randomized Controlled Trial (RCT)

► Theory-Driven Interventions

## Randomized Trial

► Experimental Design
► Factorial Design

## Rankings of Components

► Composite Indicator(s)

## Rank-Order Correlation

Michael Eid
Department of Education and Psychology,
Free University of Berlin, Berlin, Germany

### Synonyms

Correlation of ordinal variables

### Definition

A rank-order correlation is a correlation between two variables whose values are ranks.

### Description

When variables are measured at least on ordinal scales, units of observation (e.g., individuals, nations, organizations, values) can be ranked. A ranking is an ordering of units of observations with respect to an attribute of interest. For example, nations can be ranked with respect to their quality of life, their freedom, their tightness, or looseness, etc. A rank is the position of a unit of observation (e.g., nation) in the ranking. Units of observation with higher ranks show the attribute of interest to a higher degree. If one is interested in the association between two rankings (e.g., quality of life and freedom of nations), rank-order correlations can be calculated. There are many different correlation coefficients that have been developed for ordinal variables (Agresti, 1984; Gibbons & Chakraborti, 2011).

The selection of an appropriate rank-order correlation depends on whether or not there are tied observations. Two observations are tied when they share the same rank. Moreover, the choice of the correlation coefficient depends on the reasons for the observed ties. If the attribute is measured by a categorical variable with ordered categories such as a Likert scale, the ties are produced by the ► response format. There will be many units of observation that share the same response category and therefore the same rank, if units of observation are ordered with respect to the attribute. A different situation arises when the ranking is based on pair comparisons. In this case, ties are not the product of the response format but a result of the pair comparison. Each possible pair can be compared, and there are as many possible ranks as units of observation considered. If two units of observation are indistinguishable with respect to an attribute, they will share the same rank. Other examples are situations where the rankings are based on the result of a competition (e.g., ranking in a marathon run) or on continuous variables such as reaction time, speed, age, and income. In these cases, tied observations are not enforced by the response scale but a result of the comparison of individual scores that are not limited to a fixed number of single categories. Moreover, in contrast to categorical variables, there might be only few tied observations. To distinguish rankings based on categorical response variables from rankings in the other situations described above, ordinal categorical variables will be distinguished from ordinal non-categorical variables in this entry. In the following, a prototypical rank-order correlation for the following four typical data situations will be described:

1. Two ordinal non-categorical variables without tied observations
2. Two ordinal categorical variables

3. Two ordinal non-categorical variables with tied observations
4. An ordinal categorical and an ordinal non-categorical variable

Notation: In order to distinguish sample from population correlation coefficients, the sample coefficients are either marked by a hat (^) or by a Latin letter (in contrast to Greek letters for population parameters). In this entry, only the sample correlation coefficients will be described.

### Two Ordinal Non-categorical Variables Without Tied Observations

If there are no tied observations, the correlation between two ordinal non-categorical variables can be estimated by Kendall's $\tau$ (Kendall & Gibbons, 1990):

$$\hat{\tau} = \frac{n_K}{\frac{n(n-1)}{2}} - \frac{n_D}{\frac{n(n-1)}{2}} = \frac{n_K - n_D}{\frac{n(n-1)}{2}} = \frac{2\,(n_K - n_D)}{n(n-1)} \quad (1)$$

Kendall's $\tau$ depends on the comparison of all $n$ units of observation. If there are $n$ unit of observation, there are $n \cdot (n–1)/2$ possible pair comparisons (denominator in the equation). A pair comparison can show that a pair is concordant or discordant. A pair is concordant if one unit of observation has higher rank values on both variables than the other unit of observation. A pair is discordant if one unit of observation has a higher rank on one variable but a lower rank on the other variable compared to the other unit of observation. The number of concordant pairs is $n_K$, and the number of discordant pairs is $n_D$. Kendall's $\tau$ is the difference between the relative frequency of concordant pairs and the relative frequency of discordant pairs. The correlation is 0, if there are as many concordant as discordant pairs. The correlation is positive, if there are more concordant than discordant pairs. The correlation is 1 (upper limit), if there are only concordant pairs. The correlation is negative if there are more discordant pairs than concordant pairs. The correlation is $-1$ (lower limit) if there are only discordant pairs. Kendall's $\tau$ is typically applied to ordinal non-categorical variables, because tied observations cannot be avoided for ordinal categorical variables (if the sample is sufficiently large and the number of categories small).

### Two Ordinal Categorical Variables

If two units of observation are tied at least on one variable, the pair is neither concordant nor discordant. However, the tie might be due to the fact that the categories are not fine-grained enough to detect differences between the two units of observation. A more fine-grained scale might be able to detect that the pair is concordant or discordant and not tied. Because the tie might be enforced by the response format, pairs that are neither concordant nor discordant are not considered in calculating an association coefficient. The correlation between two ordinal categorical variables can be estimated by the $\gamma$-coefficient (Goodman & Kruskal, 1954):

$$\hat{\gamma} = \frac{n_K - n_D}{n_K + n_D} \quad (2)$$

The $\gamma$-coefficient contrasts the proportion of concordant pairs of all pairs that are either concordant or discordant with the proportion of discordant pairs of all pairs that are either concordant or discordant.

### Two Ordinal Non-categorical Variables with Tied Observations

If tied observations are not due to a categorical response format but the result of a pair comparison with respect to criteria that are not limited to few categories (e.g., time for a marathon run), then tied observations show that the association is weaker than it could be. If a pair has the same rank on one variable but different ranks on the other variable, then the rankings on the two variables are less similar than in the situation where one unit of observation has higher ranks on both variables. Therefore, the ties have to be considered. An appropriate coefficient for this situation is Wilson's $e$-coefficient (Wilson, 1974):

$$e = \frac{n_K - n_D}{n_K + n_D + n_{B\,(X)} + n_{B\,(Y)}} \quad (3)$$

In the denominator, $n_{B(X)}$ is the number of pairs that are tied only on variable $X$, but not on variable $Y$, whereas $n_{B(Y)}$ is the number of pairs that are tied only on variable $Y$, but not on variable $X$. If these frequencies are different from 0, the $e$-coefficients cannot take its maximum values of $-1$ or 1. Pairs with ties on both variables are not considered in Eq. 3 because they do not disturb the similarities in ranking.

### An Ordinal Categorical and an Ordinal Non-categorical Variable

If the correlation between an ordinal categorical variable and an ordinal non-categorical variable with tied observations should be calculated, the ideas of the $\gamma$- and the $e$-coefficient are combined. Such a coefficient was proposed by Somer (1962) and by Kim (1971). If the ordinal non-categorical variable is denoted by $Y$, then Somer's coefficient $d_{YX}$ is defined by

$$d_{YX} = \frac{n_K - n_D}{n_K + n_D + n_{B(Y)}} \tag{4}$$

In this equation, ties on the non-categorical variable are considered, but only when there are no ties of this pair on the categorical variable $X$ as well. If there are no tied observations on the ordinal non-categorical variable, Eq. 4 reduces to the $\gamma$-coefficient.

### Hypothesis Testing

The equations presented are equations for estimating rank-order correlations on the basis of a sample of units of observation from an underlying population. The corresponding population formula can be obtained by replacing the sample frequencies by their corresponding population probabilities. In order to test hypotheses concerning the population values as well as to calculate confidence intervals for the correlation coefficients, the sampling distributions of these coefficients have to be known. Woods (2007) reports standard errors for the coefficients considered in this entry and other correlation coefficients for rank-order data.

### Other Correlation Coefficients

Besides the coefficients presented, there are many other rank-order correlation coefficients (Agresti, 1984; Gibbons & Chakraborti, 2011; Woods, 2007). For example, Spearman's rank correlation (Spearman, 1904) is widely used in the social and behavioral sciences. Spearman's coefficient is the product–moment correlation of the rank values. Hence, it assumes that there is a linear dependency between the ranks. This strong assumption is not made by the other coefficients presented. Another well-known coefficient is the polychoric correlation that is appropriate for two ordinal categorical variables (Olsson, 1979). The polychoric correlation model assumes that each observed categorical variable is linked to an underlying continuous variable by a threshold relationship. The underlying continuous variable is divided into different categories by the threshold values. Furthermore, it is assumed that the two underlying continuous variables are distributed according to a bivariate normal distribution. The polychoric correlation usually shows very similar values to the $\gamma$-coefficient. The $\gamma$-coefficient, however, does not make this distribution assumption.

## Cross-References

- ▶ Categorical Data Analysis
- ▶ Correlation Coefficient
- ▶ Response Format

## References

Agresti, A. (1984). *Analysis of ordinal categorical data*. New York: Wiley.

Gibbons, J. D., & Chakraborti, S. (2011). *Nonparametric statistical inference* (5th ed.). Boca Raton, FL: CRC Press.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association, 49*, 732–764.

Kendall, M. G., & Gibbons, J. D. (1990). *Rank correlation methods*. London: Arnold.

Kim, J. (1971). Predictive measures of ordinal association. *The American Journal of Sociology, 76*, 891–907.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika, 44*, 443–460.

Somers, R. H. (1962). A new asymmetric measure of association for ordinal variables. *American Sociological Review, 27*, 799–811.
Spearman, C. S. (1904). The proof and measurement of association. *The American Journal of Psychology, 15*, 72–101.
Wilson, T. P. (1974). Measures of association for bivariate ordinal hypotheses. In H. M. Blalock (Ed.), *Measurement in the social science: Theories and strategies* (pp. 327–342). Chicago: Aldine Publishing.
Woods, C. M. (2007). Confidence intervals for gamma-family measures of ordinal association. *Psychological Methods, 12*, 185–204.

## Rapport Among Researchers and Subjects

► Research Relationship(s)

## Rasch Analysis

Matthias von Davier
Center for Global Assessment, Educational Testing Service, Princeton, NJ, USA

### Synonyms

Rasch model

### Definition

Rasch analysis describes procedures that use a particular model with outstanding mathematical properties developed by Georg Rasch (1960) for the analysis of data from tests and questionnaires in psychology, education, and other fields.

### Description

Georg Rasch was a Danish mathematician who developed a model for the analysis of test and attainment data (Rasch, 1960) with unique mathematical properties. This model is commonly referred to as the Rasch model but sometimes also called the 1-parameter logistic model, especially in research literature originating in North America.

The Rasch model assumes that the probability of a correct response to a certain test question – often called an item – by a human respondent depends on only two quantities, one that quantifies the difficulty of the item and one that quantifies the ability of the respondent to solve the item. This means the model makes stronger assumptions about the data than item response theory (IRT) and may therefore not be appropriate for a given set of test items. However, if the Rasch model is appropriate, it allows "specific objective comparisons": It allows comparing respondents based on the difference of person abilities without regard of the items that were involved, and differences of items in terms of their difficulty can be assessed independent of the sample of persons used to determine the difficulties.

The Rasch model can be derived mathematically from the following three assumptions:

1. The number of items that the test taker has answered correctly is a sufficient statistic of test taker's ability.
2. The probability of the test taker answering a question correctly increases with his or her ability.
3. There are no interactions among items; the response to a test item does not affect test taker's ability to correctly respond to another item.

While these assumptions are phrased in terms of cognitive ability or proficiency testing, the Rasch model is also suitable for analysis of motor skills, attitudes, and other traits. The applicability of the approach, in particular also the polytomous version of the Rasch model, is not limited to cognitive tests.

#### Rasch Analysis Specifics

The Rasch model assumes $P(X = 1|i,j) = a_j/(a_j + d_i)$, where $X$ is the response variable and $X = 1$ denotes a correct (or positive) response while $X = 0$ denotes an incorrect response and $a_j$ denotes the ability of respondent $j$ and $d_i$ denotes the difficulty of item $i$. Rasch models for

responses with more than two ordered score points are described in the entry on polytomous Rasch models. This formulation of the Rasch model originates in the papers by Rasch (1960, 1966) and shows important features of the model at a very basic level. Also, this definition is in line with and clarifies relationships of the Rasch model to models for paired comparisons (Zermelo, 1929; Bradley & Terry, 1952). Note that the ratio $a_j/(a_j + d_i)$ can be transformed into a notation that may be more familiar to an audience with experience in logistic (regression) models or IRT (Lord & Novick, 1968) by using an exponential transformation, namely, by assuming $a_j = exp(\Theta_j)$ and $d_i = exp(b_i)$, we arrive at

$$P(X = 1|i,j) = exp(\Theta_j)/(exp(\Theta_j) + exp(b_i)) = exp(\Theta_j - b_i)/(1 + exp(\Theta_j - b_i)),$$

where the right-hand side of the equation represents a notation that is commonly found in textbooks.

The simple ratio $a/(a + d)$, however, has a more intuitive appeal: Assume the item is a simple motor task, in quality of life research, this could be an everyday activity such as opening jars of preserved fruits with different diameters by hand. In sports, a task such as a free throw in basketball from a certain distance recorded in meters ($m$) could be considered. Let $a_{ji}$ denote the number of successful trials by a person $j$ while $d_{ji}$ denotes the number of unsuccessful trials, and $a_{ji} + d_{ji}$ is the total number of trials for task $i$ and person $j$. It should be obvious that the task $i$ and the person $j$ "interact"; either trying free throws from distance $m$ or opening a jar of a certain diameter requires certain motor skills. The ratio $a/(a + d)$ is simply the success rate. If $a_{ji} = d_{ji}$, then the success rate is 0.5 or 50 %. If $d_{ji} = 0$, then the success rate is 1.0 or 100 %, and if $a_{ji} = 0$, the success rate is 0.0 or 0 %. A fundamental statistical result, the *law of large numbers*, ensures that with an increasing number of trials $N = (a + d)$, the ratio $a/(a + d)$ will (in probability) get closer and closer to the expected (true) success rate, which equals the *probability of success in the long run*.

In applications of tests and questionnaires, however, tasks or questions cannot be meaningfully repeated in most cases. Each test question only appears once. The fundamental idea is to generalize across trials or tasks, instead of assuming that each task $i$ is independently attempted by a person $j$ with a unique skill represented by $a_{ji}$: We assume that person $j$ uses the same skill $a_j$ for all tasks in the test. Along the same lines, we assume that each task has only one "difficulty" $d_i$ that stands independently of the person's skill level so that we can utilize these assumption to arrive at the Rasch model: $P(X = 1|i,j) = a_j/(a_j + d_i) = exp(\Theta_j - b_i)/(1 + exp(\Theta_j - b_i))$ For details about the model, model extensions, and applications, a variety of volumes and short introductions are available; examples are Fischer and Molenaar (1995) and Andrich (1988).

## Applications to Health and Quality of Life Measures

Rasch analysis has applications in psychological testing, educational measurement, item banking, cognitive psychology, marketing, sports sciences, patient-reported outcome measurement, to name a few. The edited volume by von Davier and Carstensen (2007) provides an overview of some fields of application. For the application of Rasch analysis to quality of life measures, we list two examples here:

### WHOQOL

Power, Bullinger, Harper, and the WHOQOL group (1999) describe the development and cross-cultural validation of a World Health Organization quality of life measure (WHOQOL). The development of the WHO measures of quality of life was carried out in 15 countries, while the project has now over 40 units involved. One important aspect of the analyses of cross-cultural findings was to ensure that WHOQOL measure addresses the same construct in all country-specific versions. Among the tools chosen to confirm that this is the case was the unidimensional Rasch measurement model (Rasch, 1960, 1966) as implemented in the RUMM 2010 (Andrich, 2001) and in the WINMIRA 2001 software programs (von Davier, 2001).

### SF36 and SF12

The SF-36 (Ware, Gandek & The IQOLA Project Group, 1994) is a short health survey questionnaire with 36 items. It yields measures of functional health and well-being scores and other summary measures. The instrument has proven useful in a number of studies. Recent studies of the structure of the SF-36 have used Rasch analysis and compared the estimates with simple summated ratings (Haley, McHorney, & Ware, 1994; McHorney, Haley, & Ware, 1997). Among the practical implications of the Rasch analysis, authors mention greater score precision and reduced clustering of scores at extremes of the scale, which is due to the fact that scores on IRT and Rasch scales are nonlinearly but strict monotonic related to the summated scores. More recent studies such as Raczek et al. (1998), as well as Taylor and McPherson (2007), have been using the Rasch model to analyze data from the SF-36 collected in applications to study a variety of populations.

### HAQ-II

Wolfe, Michaud, and Pincus (2004) presented a revised version of the Health Assessment Questionnaire (HAQ). The authors refer to the HAQ as the most important and widely used functional status questionnaire in rheumatology. The HAQ-II consists of 10 items and was developed using Rasch analysis using a database of more than 10,000 respondents. The original HAQ served as the basis for these developments. The authors report that by selecting items that do show good fit to the sufficiency assumptions underlying Rasch analysis and that provide good coverage of the domain, the HAQ-II emerged with improved psychometric scale properties such as increased reliability compared to the original HAQ.

## Some Selected Tools for Rasch Analysis

A variety of software programs for Rasch analysis are available. Some specialized software packages such as WINMIRA 2001 (von Davier, 2000), Rumm 2010 (Andrich, 1999) and Conquest (Wu, Adams, & Wilson, 1997), to name a few, are commercially available. While RUMM is based on an estimation method that is less commonly applied, WINMIRA 2001 and Conquest allow conditional maximum likelihood estimation of item parameters and subsequent estimation of person parameters. There are some older software packages that should not be used anymore since these are using estimation methods (more specifically, joint maximum likelihood methods) that have been shown to produce biased estimates (Andersen, 1972; Haberman, 1977). The full range of software packages with a detailed discussion of advantages and disadvantages cannot be covered here. In addition, there are add-on packages capable of Rasch analysis for use with general-purpose statistical software such as R, STATA, and SAS. As an example, the eRm package (Mair & Hatzinger, 2007) can be used within R to obtain conditional maximum likelihood estimates. As a general rule, these add-on packages will be less capable in terms of run-time speed, maximum sample size, and numbers of tasks/items in the data since these packages are interpreted and executed under a general-purpose tool and not available in the form of run-time-optimized code.

## Cross-References

▶ Item Response Theory [IRT]
▶ Rasch Polytomous Models

## References

Andersen, E. B. (1972). The numerical solution of a set of conditional estimation equations. *Journal of the Royal Statistical Society, Series B, 34*, 42–54.

Andrich, D. A. (1988). *Rasch models for measurement* (Quantitative methods in social sciences, Vol. 68). Newbury Park: Sage.

Andrich, D. (2001). *RUMM2010: Rasch unidimensional measurement models*. Perth: RUMM Laboratory.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs. I. the method of paired comparisons. *Biometrika, 39*, 324–345.

Fischer, G. H., & Molenaar, I. W. (1995). *Rasch models – Foundations, recent developments and applications*. New York: Springer.

Haberman, S. J. (1977). Maximum likelihood estimates in exponential response models. *The Annals of Statistics, 5*, 815–841.

Haley, S. M., McHorney, C. A., & Ware, J. E. (1994). Evaluation of the MOS SF 36 physical functioning scale (PF 10): I. Unidimensionality and reproducibility of the Rasch item scale. *Journal of Clinical Epidemiology, 47*(6), 671–684.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Mair, P., & Hatzinger, R. (2007). Extended Rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software, 20*(9), 1–20.

McHorney, C. A., Haley, S. M., & Ware, J. E. (1997). Evaluation of the MOS SF-36 physical functioning scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *Journal of Clinical Epidemiology, 50*(4), 451–461.

Power, M. J., Bullinger, M., Harper, A., & The WHOQOL Group. (1999). The World Health Organization WHOQOL-100: Tests of the universality of quality of life in fifteen different cultural groups worldwide. *Health Psychology, 18*, 495–505.

Raczek, A. E., Ware, J. E., Bjorner, J. B., Gandek, B., Haley, S. N., Aaronson, N. K., et al. (1998). Comparison of Rasch and Summated rating scales constructed from SF-36 physical functioning items in seven countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology, 51*(11), 1203–1214.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Rasch, G. (1966). An item analysis that takes individual differences into account. *British Journal of Mathematical and Statistical Psychology, 19*(1), 49–57.

Taylor, W., & McPherson, K. (2007). Using Rasch analysis to compare the psychometric properties of the Short Form 36 physical function score and the Health Assessment Questionnaire disability index in patients with psoriatic arthritis and rheumatoid arthritis. *Arthritis & Rheumatism*, *57*(5), 723–729. WILEY-LISS. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17530670

von Davier, M. (2001) WINMIRA 2001. A windows-program for analyses with the Rasch model, with the latent class analysis and with the mixed Rasch model. Software distributed by ASC – Assessment Systems Corporation (USA) and iec ProGAMMA (The Netherlands).

von Davier, M., & Carstensen, C. H. (2007). *Multivariate and mixture distribution Rasch models – Extensions and applications*. New York: Springer. ISBN 978-0387-32916-1.

Ware, J. E., Gandek, B., & The IQOLA Project Group. (1994). The SF-36® Health Survey: Development and use in mental health research and the IQOLA project. *International Journal of Mental Health, 23*(2), 49–73.

Wolfe, F., Michaud, K., & Pincus, T. (2004). Development and validation of the health assessment questionnaire II: a revised version of the health assessment questionnaire. *Arthritis and Rheumatism, 50*(10), 3296–3305.

Wu, M. L., Adams, R. J., & Wilson, M. R. (1997). *ConQuest: Multi-aspect test software, [computer program]*. Camberwell: Australian Council for Educational Research.

Zermelo, E. (1929). Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. (The calculation of tournament results as a maximum problem of probability calculus: in German). *Mathematische Zeitschrift, 29*, 436–460.

## Rasch Analysis and Item Response Theory (IRT)

▶ Item Response Theory [IRT]

## Rasch Model

▶ Rasch Analysis

## Rasch Polytomous Models

Matthias von Davier
Center for Global Assessment, Educational Testing Service, Princeton, NJ, USA

### Synonyms

Partial credit model; Polytomous Rasch model; Rating scale model

### Definition

Rasch polytomous models are statistical models for test and questionnaire data suitable for the analysis of data collected using rating scales, Likert-type response scales, or other response data with ordered categories, while preserving the main defining characteristics of ▶ Rasch analysis for binary responses.

## Description

The ▶ Rasch model is often initially introduced as a model for binary data. In that case, we consider 0 (zero) as an indicator of a negative or incorrect response, while we consider 1 (one) as indicating a positive or correct response. In questionnaires used to assess quality of life, patient reported outcomes, or personality traits, however, we often find a response format that is more fine-grained.

The ▶ Likert scale (Likert, 1932) is maybe the most commonly used of these ordinal response formats: The responses in this rating scale are ordered, ranging from a low end that typically represents complete disagreement with the statement to an upper end that represents complete agreement with the statement. Respondents "rate" their level of agreement to each of the statements in a questionnaire using the same verbal or numerical anchors. Rasch models for these ordered response formats are often direct generalizations of binary Rasch models described in the volume's entry on Rasch analysis.

The fact that multiple responses are available per item while only one can be chosen by the respondent has certain consequences for the formalization of models that are consistent with the basic assumptions of Rasch analysis. As an example, the binary Rasch model has a simple sufficient statistic, the total score, which is all that is required for estimating respondent's ability; this is a feature that should be retained by Rasch models for polytomous ordinal response formats.

The models that fulfill this requirement have been developed over time, and estimation methods for a number of these polytomous Rasch models have been described in the research literature (Andrich, 1978; Rost, 1991; von Davier & Rost, 1995). While all models in this entry assume the responses to have a finite number of ordered choices, there are models that are consistent with the basic assumptions of Rasch analysis but can be used for continuous responses (Müller, 1987) or are suitable for count data (e.g., Jansen, 1995). However, these will not be discussed in detail here since they are less commonly used for quality of life and other self-report questionnaire data.

Typically, the responses to Likert-type ordinal response formats are numerically represented by successive integers. To be consistent with binary data where the positive response is coded as 1 and the negative is represented by 0, Likert item responses are often coded from 0, 1, 2, 3, and 4, for a response format with m + 1 categories, obtaining a set of responses that can be represented as the integers from 0 to m.

It is convenient to write the Rasch model for polytomous response data in the form of a general model equation first and then to introduce the different variants introduced by Andrich, Masters, and Rost. The general form can be written as

$$P(X = x|i,j) == \exp(\alpha_i x_j) / (\Sigma x \exp(\alpha_i x_j)),$$

with $\Sigma x \exp(\alpha_i x_j) = \exp(\alpha_i 0_j) + \exp(\alpha_i 1_j) + \ldots + \exp(\alpha_i m_j)$ necessary for sum normalization. For the following exposition, we assume without loss of generality that $\alpha_i 0_j = 0$ for the lowest category, setting this response as the reference, and then define the model by specifying the $\alpha_i x_j$ for the responses $x = 1, \ldots, m$. Then, the partial credit model (Masters, 1982) is defined by

$$\alpha_i x_j = x\theta_j - \beta_{ix}$$

and is probably the most commonly used among the polytomous ordinal Rasch models. Just as in the binary Rasch model, the parameter $\theta_j$ denotes the person's ability or attitude level. In contrast to the binary Rasch model, multiple-item parameters, sometimes called (cumulative) threshold parameters, $\beta_{ix}$ are required, one for each transition between a lower category $x - 1$ to category $x$. These parameters quantify the "difficulty" of transition from lower to higher response categories.

There are various ways in which this model can be modified to accommodate certain assumptions about the order structure of the response format. Andrich (1978, 1982) suggests the following two constraints imposed on the parameters. One constraint, referred to as the rating scale model in the following, assumes that the "distance" between the categories of the rating scale is the same across items. The other approach,

referred to as the equidistance model here, assumes that the "distance" between categories of the rating scale is constant for each item. The rating scale model is given by

$$\alpha_i x_j = x\theta_j - (\sigma_i + \tau_x)$$

and has one category parameter for each of the responses $x = 1, \ldots, m$. Note that this parameter $\tau_x$ is the same across all items, only the average difficulty σi can be different for different items $i$. The equidistance model is given by

$$\alpha_i x_j = x\theta_j - (\sigma_i + [m + 1 - 2x]\delta_i)$$

and has one parameter $\delta_i$ that can differ from item to item. Rost (1988) combined the two approaches and suggested the dispersion model

$$\alpha_i x_j = x\theta_j - (\sigma_i + \tau_x + [m + 1 - 2x]\delta_i).$$

One of the advantages of these special cases of the partial credit model is that they allow a more parsimonious representation of the model. For example, the partial credit model for a test with 25 items with a 5-category response format would require statistical estimation of 100 distinct threshold parameters, whereas the rating scale model would only require 25 + 4 = 29 parameters (assuming a model that centers the ability distribution around zero). In addition, these constrained models allow testing whether items are mainly differing with respect to their overall difficulty $\sigma_i$ rather than their thresholds, for example, whether it can be assumed that the following decomposition holds: $\beta_{ix} = (\sigma_i + \tau_x)$. That is, whether the rating scale model or the partial credit model should be adopted is of course not only a matter of how many parameters has to be estimated but also a question of model-data fit (von Davier & Rost, 1995).

Details on the estimation of these models and the differences between the approaches can be found in the volumes by Fischer and Molenaar (1995) and von Davier and Carstensen (2007). The WINMIRA software (von Davier, 2000) can be used to estimate model parameters for the different polytomous (ordinal) Rasch models presented here. More specifically, the "rating scale model" (Andrich, 1978), the "equidistance model" (Andrich, 1982), the "dispersion model" (Rost, 1988), and the "partial credit model" (Masters, 1982) as well as the discrete mixture distribution Rasch model generalizations of these polytomous Rasch models (Rost, 1991; von Davier & Rost; 1995) can be estimated with this software.

Adams, Wilson, and Wang (1997) propose the Mixed-Coefficient Multinomial Logit Model (MCMLM), a generalization of the Rasch model that can handle polytomous response data as well. This approach is a direct extension of the Rasch model for ordinal data for response variable that measure multiple, potentially different personality characteristics. Moreover, each response category can be weighted differently (with a priori fixed weights). Adams, Wu, and Carstensen (2007) report the application of this model in the analysis of international large-scale assessment data. A commercial software (ConQuest) can be used to estimate the MCMLM.

The general diagnostic model (GDM; von Davier, 2005) is also suitable for uni- and multidimensional Rasch analysis and for analysis with the polytomous Rasch model (Xu, 2007). The GDM approach represents a whole family of models and allows, for example, a straightforward extension of the partial credit model (Masters, 1982) and generalizations (Muraki, 1992) to several latent dimensions with discrete and quasi-continuous structure. von Davier (2010) extends the GDM approach to a nonparametric multilevel IRT model (including Rasch models) for ordinal data.

## References

Adams, R. J., Wilson, M. R., & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*, 1–23.

Adams, R. J., Wu, M. L., & Carstensen, C. H. (2007). Application of multivariate Rasch models in international large scale educational assessment. In M. von Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models: Extensions and applications* (pp. 271–280). New York: Springer.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561–573.
Andrich, D. (1982). An extension of the Rasch model for ratings providing both location and dispersion parameters. *Psychometrika, 47*, 105–113.
Fischer, G. H., & Molenaar, I. W. (1995). *Rasch models – Foundations, recent developments and applications*. New York: Springer.
Jansen, M. G. H. (1995). The Rasch poisson counts model for incomplete data: An application of the EM algorithm. *Applied Psychological Measurement, 19*(3), 291–302.
Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 1–55.
Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 147–174.
Müller, H. (1987). A Rasch model for continuous ratings. *Psychometrika, 52*, 165–181.
Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*, 159–176.
Rost, J. (1988). Measuring attitudes with a threshold model drawing on a traditional scaling concept. *Applied Psychological Measurement, 12*, 397–409.
Rost, J. (1991). A logistic mixture distribution model for polytomous item responses. *British Journal of Mathematical and Statistical Psychology, 44*, 75–92.
von Davier, M. (2000). WINMIRA 2001 - Software Manual. IPN: *Institute for Science Education*. Student Version available at: http://www.von-davier.com
von Davier, M. (2005). *A general diagnostic model applied to language testing data* (ETS Research Report RR-05-16). Princeton, NJ: Educational Testing Service.
von Davier, M. (2010). Hierarchical mixtures of diagnostic models. *Psychological Test and Assessment Modeling*, *52*(1), 8–28. Retrieved May 10, 2010, from http://www.psychologie-aktuell.com/fileadmin/download/ptam/1-2010/02_vonDavier.pdf
von Davier, M., & Carstensen, C. H. (Eds.). (2007). *Multivariate and mixture distribution Rasch models: Extensions and applications*. New York: Springer.
von Davier, M., & Rost, J. (1995). Polytomous mixed Rasch models. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch models – Foundations, recent developments and applications* (pp. 371–379). New York: Springer.
Xu, X. (2007). *Monotone properties of a general diagnostic model* (ETS Research Rep. No. RR-07-25). Princeton, NJ: Educational Testing Service.

## Rat Race of Consumption

► Consumption Externalities

## Rating Form of IBD Patient Concerns (RFIPC)

► Health-Related Quality of Life and Inflammatory Bowel Disease

## Rating Scale

► Thermometer Scale (Feeling Thermometer)

## Rating Scale Model

► Rasch Polytomous Models

## Ratings of Components

► Composite Indicator(s)

## Ratio Measurement

► Ratio Scales

## Ratio Response Format

► Response Format

## Ratio Scales

Ester Cerin
Institute of Human Performance, The University of Hong Kong, Hong Kong, People's Republic of China

### Synonyms

Ratio measurement; Ratio variables

## Definition

Ratio scales are measurement scales that consist of equidistant units and have a nonarbitrary zero point, signifying the absence of a measured characteristic (Field, 2009; Michell, 1999; Stevens, 1946).

## Description

Ratio scales are one of the four ▶ levels of measurement which describe the relationship between what is being measured and the values assigned to what is being measured. They represent the highest level of measurement within Stevens' (1946) theory of scale types. They are rare in the social sciences but common in the physical sciences and engineering. Examples of ratio scales are age, income, length, time, and mass. Within ▶ quality of life research, ratio scales usually represent independent (e.g., age, income, time spent engaging in physical activity) rather than dependent variables (▶ SF-36 Health Survey).

Ratio scales allow four basic empirical operations including classification (e.g., determination of equality), rank ordering, determination of equality of intervals or differences, and determination of equality of ratios (Stevens, 1946). Determination of equality refers to the fact that when using ratio scales, two objects with the same amount of a variable are assigned the same value, while objects with different amounts of a variable are signed different values. For example, the length of a 30-cm-long stick is different from that of a 15-cm-long stick. Rank ordering pertains to the ability to determine whether a quantity is greater, less, or equal to another quantity. Using the example above, we know that a 30-cm stick is longer than a 15-cm stick. Determination of equality of intervals or differences refers to the fact that the points on a ratio scale are equidistant. Thus, the difference in length between a 30- and a 15-cm stick is the same as the difference between a 24- and a 9-cm stick. In other words, an x-unit difference along any section of the scale always has the same meaning. Finally, the determination of equality of ratios is the basic empirical operation that distinguishes ratio scales from lower levels of measurement. The ratios of the values along a ratio scale are meaningful. This is because this type of scale has a nonarbitrary, true zero. For example, we can state that the length of a 30-cm stick is twice the length of a 15-cm stick, since 30 divided by 15 is 2. The same applies to a 100-cm stick as compared to a 50-cm stick as the ratio of these values is also 2.

Ratio scales allow the use of all available measures of ▶ central tendency and measures of dispersion to describe a set of data (Field, 2009; Stevens, 1946). These include the mode, median, arithmetic mean, geometric mean, harmonic mean, range, interquartile range, standard deviation, and coefficient of variation.

## Cross-References

▶ Central Tendency
▶ Levels of Measurement
▶ Quality of Life
▶ SF-36 Health Survey

## References

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.
Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*. Cambridge, UK: Cambridge University Press.
Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.

# Ratio Variables

▶ Ratio Scales

# Rational Behavior

▶ Economic Rationality Assumption

## Rational Choice

► Economic Rationality Assumption

## Rational Choice Theory

Bereket Kebede
School of International Development (DEV), University of East Anglia, Norwich, UK

### Synonyms

Optimal choice theory; Optimization theory; Rational decision-making; Rationality

### Definition

Rational choice theory is a formal choice theory that starts from arguably widely accepted basic axioms and builds a logically consistent decision model for maximizing the objective of a decision-maker. Formal rational choice theory was mainly developed and used in economics, particularly microeconomics, but its influence reaches far wider than economics.

### Description

Rational choice theory is usually identified with the decision theory mainly developed within economics; the theory is intimately related to the rise and formalization of "neo-classical" economics. Rational choice theory focuses on the choice of individuals and hence is an advocate of "methodological individualism" – even higher level societal choices are grounded in the decisions of individuals. The formal theory starts from relatively simple but arguably widely accepted assumptions about preferences of individuals and then builds a formal decision-making framework where net benefits (benefits minus costs) of individuals are maximized. The preferences of individuals are assumed to be complete. For example, in consumption the individual is assumed to have stable preference over all the goods in the relevant consumption bundle either in the form of preferring one good over the other or being indifferent; in other words, all alternatives are comparable. Preferences are also assumed to be transitive implying that if the individual prefers commodity $x$ to $y$ and prefers commodity $y$ to $z$, then $x$ would be preferred to $z$ (Varian 1992). Given a stable preference that satisfies the underlining assumptions, the decision problem is maximizing net benefits constrained by available resources. This constrained maximization problem invariably leads to a decision rule where marginal net benefits should be equalized. For example, in consumption to maximize satisfaction (utility), the marginal utility per unit of money should be equal across all goods consumed. If not, the consumer will still have the opportunity to increase utility by reallocating money from goods that give less marginal utility to those that give more marginal utility per unit of money.

How the behavior of others affects the choice of the individual and vice versa has to be considered when the decision scenario expands to more than one individual. Game theory was developed to understand this. In game theory, the strategic interaction of rational decision-makers is modeled, and equilibrium concepts (like Nash equilibrium) are developed to understand likely outcomes.

Even though rational choice theory is dominant in conventional economics, its influence has a much wider reach. In sociology, in contrast to approaches that emphasize the importance of "structure" (the social environment) on individual decision-making, rational choice theory gives priority to "agency" where individuals are considered as conscious decision-makers calculating the costs and benefits of different alternatives (Hedström and Stern 2008). Rational choice theory has also strongly influenced psychology, political science, and political philosophy (Oppenheimer 2008). The wider influence of rational choice theory is evidenced by the popularity of game theory in many social science disciplines.

The growing importance of rational choice theory does not mean that it is immune from criticisms. Rational choice theory has been criticized from very different angles. The critical literature is voluminous, and this short piece can only highlight a few texts. A strong criticism of the formal structure of rational choice theory is provided by Amartya Sen (Sen 1977; Sen 1982; Sen 1994) among others. Part of his criticism is directed toward the lack of clarity about the nature of preferences in the theory; it is not clear whether preferences capture satisfaction or pleasure, introspective welfare, desired choice, normative judgement, or other aspects. He coined the term "rational fool" for a decision-maker who is rigorously consistent, as formally defined in the theory, but does not make clear distinctions between the different aspects of preference. The link between choice and individual well-being may be broken if behavior is influenced by motivations like commitment and morality. For example, one can be committed to remove some injustice even though one does not suffer from it, and the commitment can be costly for the individual (Sen 1977). Interdependence between decision-makers may lead to welfare outcomes that are inferior, as in the prisoners' dilemma case, even though the choices of individuals separately are optimal (Sen 1982).

Some criticize rational choice theory for its focus on individual decision-making and for neglecting social interactions. Since social interaction is the basis of human life, rational decision theory is considered unacceptable, especially in sociology. A focus on social networks through which individuals interact and learn is called for (Pescosolido 1992). Even though individual decision-making is the starting point for rational choice theory, this criticism is not completely fair if the many extensions of the theory to analyze social interactions are considered. In economics, Gary Becker is one of the prominent forerunners who showed that rational choice theory can be used to understand many social issues. By incorporating richer attitudes and preferences, the conventional rational choice theory has been fruitfully extended to analyze social issues like discrimination against minorities, crime and punishment and formation, and dissolution and structure of families (Becker 1993). Following the pioneering work of Becker, a huge literature has developed in this area.

Another group of criticisms focus on the assumption that decision-makers are purely selfish (egoistic). Even though self-interest definitely plays an important role in decisions of individuals, a lot of research, mainly in behavioral and experimental economics, has shown that social preferences like fairness, trust, reciprocity, and altruism also significantly influence decisions. Different experimental games (like the dictator, ultimatum, trust, and similar games in their myriad versions) have shown that people care about fairness, they are willing to pay a price to redress what they perceive as unfair, and their decisions are significantly influenced by "social image" (how another person perceives the behavior of an individual). Interestingly, when competition is introduced in these experimental games, generally the behavior of players approaches that of the selfish individual. These results imply that the selfish decision-maker assumed in rational choice theory is probably a good characterization of behavior when competitive pressures increase; competition seems to crowd out other-regarding behavior (Camerer 2010).

Some criticisms relate to the capacity for optimization/maximization. In rational choice theory, individuals are assumed to have the capacity to solve relatively complicated maximization problems – they are "hyperrational." Solving an optimization problem requires mental capacity, and obviously the mental capacity of people differs. In addition, solving a maximization problem requires effort; people have to think hard, process information, exert their memory power, etc. In the conventional rational choice theory, both heterogeneity in mental capacity of individuals and the cost of solving optimization problems are ignored. In standard game theory, decision-makers are assumed to have the capacity for infinite iterative reasoning. All decision-makers are assumed to work out the effect of their actions on others, effect of these reactions on their decisions, etc. That is, they can work out the implication of the infinite sequence of actions and reactions. This obviously

requires a lot of mental effort, and most people do not have the capacity to do that. Bounded rational models consider these limits in human capacity and analyze the implications on solving complicated optimization problems (e.g., see Mullainathan 2002). Many experimental games suggest that most people use only between 0 and 2 steps in their iterative reasoning (Camerer 2010). A fast-growing literature is building around the implication of this heterogeneity in the capacity of decision-makers. Here, not only how many steps of iterative reasoning an individual considers but also the individual's expectation about how many steps others think through are important determinants of behavior (Camerer 2010).

Even though criticized from different angles and limited in many respects, rational choice theory is still highly influential not only in economics but also in other disciplines. In addition to its own worth as a useful tool of analyzing decision scenarios, one expects its influence to continue in the future because of other reasons. First, rational choice theory is a very good starting point to many more complicated models that reflect heterogeneity of individuals in terms of decision-making capacity. Second, as many studies have already shown, the theory can be expanded to consider a wider scope of preferences, including social preferences.

## Cross-References

- ▶ Choice
- ▶ Decision Making
- ▶ Rationality

## References


Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy, 101*(3).

Camerer, C. F. (2010). Behavioural game theory. In S. N. Durlauf & L. E. Blume (Eds.), *Behavioural and experimental economics*. Houndmills: Palgrave Macmillan.

Hedström, P., & Stern, C. (2008). Rational choice and sociology. In L. Blume & S. Durlauf (Eds.), *The New Palgrave dictionary of economics*. Basingstoke: Palgrave Macmillan.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *Quarterly Journal of Economics, 117*(3), 735–774.

Oppenheimer, J. A. (2008). Rational choice theory. In R. Adcock (Ed.), *The sage encyclopedia of political theory*. London: Sage.

Pescosolido, B. A. (1992). Beyond rational choice: The social dynamics of how people seek help. *The American Journal of Sociology, 97*(4), 1096–1138.

Sen, A. (1982). *Choice, welfare and measurement*. Cambridge, MA/London: Harvard University Press.

Sen, A. (1994). The formulation of rational choice. *American Economic Review, 84*(2), 385–390.

Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy & Public Affairs, 6*(4), 317–344.

Varian, H. R. (1992). *Microeconomic analysis*. New York/London: W.W. Norton and Company.


## Rational Decision-Making

- ▶ Rational Choice Theory

## Rationalism

- ▶ Economic Rationality Assumption

## Rationality

- ▶ Rational Choice Theory
- ▶ Economic Rationality Assumption

## Rawls' Basic Goods

- ▶ Primary Goods

## Reading Components of Literacy Assessments

- ▶ International Literacy Assessments

# Reading Literacy Achievement

Luisa Araujo
Joint Research Centre, European Commission, Ispra, Italy
Instituto Superior de Educação e Ciências, Lisbon, Portugal

## Synonyms

Functional literacy; Literacy achievement

## Definition

Reading literacy achievement is a general term used to describe the levels of reading comprehension of the school population. It is associated with large-scale assessments of reading at the fourth grade level PIRLS – Program for International Reading Literacy Study and with PISA – Program for International Student Assessment, which samples 15 year olds in OECD and partner countries. Reading achievement in both surveys can be interpreted against an international mean set at 500 and a standard deviation of 100.

## Description

Reading literacy achievement is a measure of text-based reading comprehension. Surveys such as PIRLS, conducted by the International Association for the Evaluation of Educational Achievement (IEA), assume that by fourth grade students have acquired the basic reading skills necessary to comprehend what they read. Reading comprehension is ..."an active and complex process that involves understanding written text; developing and interpreting meaning; and using meaning as appropriate to type of text, purpose and situation" (National Assessment Governing Board, 2007).

Designed to assess reading at the fourth grade level on a 5-year cycle since 2001, PIRLS seeks to capture students' performance at the beginning of the second reading stage proposed by Jeanne Chall – the *reading to learn* stage (Chall, 1996). While in the first ▶ learning *to read* stage in an alphabetical language, students focus on decoding written symbols rapidly and accurately with basis on phonological processing, they subsequently begin to read to learn by focusing their attention on interpreting meaning from print (Catts, 2009). In PISA, the Program for International Student Assessment launched by the Organization for Economic Cooperation and Development (OECD) in 2000, reading literacy for 15 year olds is conceptualized as a measure of students' ability to participate in society and further their learning to later enter the world of work (OECD, 2004).

The first survey, PIRLS, defines reading literacy as "the ability to understand and use those written language forms required by society and/or valued by the individual (Mullis, Kennedy, Martin, & Sainsbury, 2006)." PISA's definition of literacy refers to comparable goals: "the ability to understand, reflect on and use written texts, in order to achieve one's goals and to participate effectively in society" (OECD, 2000). In PISA, this ability is captured in a range of performance levels, referred to as proficiency levels that are determined using ▶ Item Response Theory (IRT) statistical models (Embretson & Reise, 2000). The statistical procedures associated with these models make it possible to assess the difficulty of a test's item and to establish specific ▶ performance indicators linked to subjects' ability levels. Performance at level 1 and below is described as insufficient to understand the simplest reading tasks and to function in society. In contrast, proficiency levels 5 and 6 reflect the ability to perform reading tasks that require a critical evaluation of information that may not be explicitly stated in a long and detailed text (OECD, 2007).

In order to capture the relationship between reading achievement and potential socioeconomic, ▶ human capital, attitudinal, and behavioral factors, both PIRLS and PISA collect data on home literacy practices. Findings from PIRLS, for example, show that fourth grade students' achievement is positively related

to home literacy activities before school starts (Mullis, Martin, Kennedy, & Foy, 2007). Similarly, results from PISA indicate that students who read more outside of school attain higher reading proficiency levels (OECD, 2010a). These findings are in line with current reading research showing that young children's vocabulary knowledge, acquired from parental book reading prior to school entry, positively impacts their future reading in grade three (Sénéchal, Ouellette, & Rodney, 2006). Conclusions from the National Reading Panel ▶ meta-analysis at the elementary school level also corroborate the PISA results, indicating that recreational reading is a privileged way to learn reading skills implicitly. Students with strong reading habits consolidate orthographic representations of words they repeatedly encounter in print and constantly learn new vocabulary (Naggy & Scott, 2000). Teacher-guided reading experiences can motivate students to engage in reading for pleasure and thus to become better readers (Guthrie, Wigfield, & Von Secker, 2000).

Clearly, these surveys provide participating countries important information about the reading literacy performance of their students. In the last 2006 PIRLS study, 45 countries participated and in 2011, 55 countries participated in PIRLS. PISA registered the participation of 43 countries in 2000 and counted with the participation of 66 countries in 2009. For many countries, trend information is available because they have participated in consecutive assessment rounds. Thus, individual countries can compare their performance to that of other countries, but also with their own performance over time. Moreover, they can assess how specific groups of students, such as immigrants, are falling behind or improving when compared to native students and whether the observed gender differences in reading are widening or narrowing.

Information gathered in PISA also serves to guide policy at the European level. For instance, the European Commission uses the percentage of low achievers in reading as an indicator (indicators, optimal level of detail) to monitor the attainment of educational objectives. The low achievers are those students performing at level 1 and below and the policy target for 2020 is that no more than 15 % of students should fall under this category in the EU-27 (DG Education and Culture Management Plan, 2011).

## Discussion

The use of international surveys to measure reading ▶ literacy achievement is congruent with the current focus on accountability movements in education that seek to monitor progress though the establishment of performance benchmarks (Ravitch, 2010). It also fits well with current perspectives linking educational achievement to ▶ economic development. The same entity that develops and implements the PISA study, OECD, has recently published a study on the high cost of low educational performance, showing that increasing students' performance in all areas tested in PISA – reading, mathematics, and science – would result in a considerable increase in a country's economic growth (OECD, 2010b).

Such clear-cut links between literacy attainment and economic development may still need further validation. However, what decades of reading research have clearly shown is that failing to learn to read fluently with good comprehension before the third or fourth year of schooling is likely to result in life-long problems in learning new skills (e.g., life-long learning). As children continue their schooling, those with poor reading skills read less and fall further and further behind. However, all through the middle and high school grades (5–12 grades), good reading skills are essential to understand content in subjects such as science and history. Thus, reading difficulties are a likely indicator of difficulty in most other school subjects and may negatively impact ▶ high school completion rates (Adams, 2009). In sum, poor reading achievement is likely to contribute to poor academic achievement which in turn may substantially reduce one's chances of completing secondary and tertiary studies and consequently impact future labor market integration and one's ▶ quality of life.

## Cross-References

▶ Economic Development
▶ High School Completion Rates

▶ Human Capital
▶ Immigrants, an Overview
▶ Item Response Theory
▶ Learning
▶ Literacy
▶ Meta-Analysis
▶ Performance Indicators
▶ Quality of Life

## References

Adams, M. J. (2009). The challenge of advanced texts: The interdependence of reading and learning. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 163–189). New York: Guilford Press.

Catts, H. W. (2009). The narrow view of reading promotes a broad view of comprehension. *Language, Speech, and Hearing Services in Schools, 40*, 178–183.

Chall, J. S. (1996). *Learning to read: The great debate (revised)*. New York: McGraw-Hill.

DG Education and Culture Management Plan. (2011). *Resource document. Education and culture directorate general*. Accessed March 15, 2011, from http://ec.europa.eu/atwork/synthesis/amp/doc/eac_mp.pdf

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: LEA.

Guthrie, J. T., Wigfield, A., & Von Secker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology, 92*(2), 331–341.

Mullis, I. V. S., Kennedy, A. M., Martin, M. O., & Sainsbury, M. (2006). *PIRLS 2006 assessment framework and specification* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *IEA's progress in international reading literacy study in primary school in 40 countries*. Chestnut Hill, MA: Boston College.

Naggy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). Mahwah, NJ: Lawrence Erlbaum.

National Assessment Governing Board. (2007). *Reading framework for the 2009 national assessment of educational progress*. Washington DC: US Department of Education.

OECD. (2000). *Knowledge and skills for life: First results from PISA 2000*. Paris: OECD Publications.

OECD. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: OECD Publications.

OECD. (2007). *PISA 2006: Science competencies for tomorrow's world executive summary*. Paris: OECD Publications.

OECD. (2010a). *PISA 2009 results: What students know and can do*. Paris: OECD Publications.

OECD. (2010b). *The high cost of low educational performance*. Paris: OECD Publications.

Ravitch, D. (2010). *The death and life of the great American school system*. New York: Basic Books.

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In *Handbook of early literacy research* (pp. 173–184). New York: Guilford Press.

# Reading Newspaper Articles

Elisabetta Lazzaro
Division of Cultural Management, Université Libre de Bruxelles (ULB), Brussels, Belgium

## Synonyms

Advertising; Articles; Information; Media use; News

## Definition

Looking at daily or weekly publications, news, editorial or feature articles, and advertising

## Description

This human activity consists in looking at, examining, and grasping the meaning of a variety of written and/or illustrated news, editorial or feature articles, as well as advertising, written by professional and other journalists and editors and which are contained in usually daily or weekly publications, available on printed support and/or on the Internet. Depending on the nature of the publication, possible reasons – and hence classifying categories (Shelley, 2005) – of reading newspapers include personal interest in general or specialized (e.g., financial, politics, arts) information, leisure and entertainment, relaxing, socializing, and knowledge and learning.

Hofstetter and Madjar (2003) provide an interesting literature review of time spent on newspaper reading and other activities and happiness. In particular, men and women's time spent in reading newspapers can be associated with more or less enjoyment and happiness, similarly to other forms of entertainment, playing sports and social activities (Graef, McManama Gianinno, & Csikszentmihalyi, 1981; Gershuny & Halpin, 1996; Robinson & Martin, 2008). Furthermore, Internet increases women's time spent on newspaper reading (Gershuny, 2002). Through an empirical study conducted among Dutch readers, Wolswinkel (2008) found new young readers prefer on-line free newspapers to printed ones. In addition, easy conditions, rather than content of reading, favor their allocated time.

## Cross-References

- ▶ Community
- ▶ Cultural Diversity
- ▶ Education
- ▶ Happiness
- ▶ Quality of Life
- ▶ Subjective Well-being
- ▶ Time Use(S)

## References


Gershuny, J. (2002). *Web-use and Net-nerds: A Neo-Functionalist Analysis of the Impact of Information Technology in the Home*. ISER Working Paper 2002–1. Colchester: University of Essex. Retrived date Nov 20, 2012 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.202.313&rep=rep1&type=pdf.

Gershuny, J., & Halpin, B. (1996). Time use, quality of life and process benefits. In A. Offer (Ed.), *In pursuit of the quality of life*. New York: Oxford University Press.

Graef, R., McManama Gianinno, S., & Csikszentmihalyi, M. (1981). Energy consumption in leisure and perceived happiness. In J. D. Claxton, C. D. Anderson, J. R. B. Richie, & G. H. G. Mc Dougall (Eds.), *Consumers and energy conversation: International Perspectives on research and Policy Options*. New York: Praeger.

Hofstetter, P. & Madjar, M. (2003). *Linking change in happiness, time-use, sustainable consumption, and environmental impacts; An attempt to understand time-rebound effects*. Final Report. Date Retrived Nov 20, 2012 from http://www.wsis.ethz.ch/hofstettermadjar.pdf.

Robinson, J. P. & Martin, S. (2008). What Do Happy People Do?. *Social Indicators Research*, *89*, 565–571. http://download.springer.com/static/pdf/675/art%253A10.1007%252Fs11205-008-9296-6.pdf?auth66=1354578851_52fc0dc7a47a63cd10b1ab61b1e14845&ext=.pdf.

Shelley, K. J. (2005). Developing the American Time Use Survey activity classification system. *Monthly Labor Review*, *128*, 3–15. http://www.bls.gov/opub/mlr/2005/06/art1full.pdf.

Wolswinkel, H. (2008). *New paper reading: The explaining factors of newspaper reading behavior of young adults*. University of Twente, Master Thesis. http://essay.utwente.nl/59108/1/scriptie_H_Wolswinkel.pdf.


## Real Choice Experiments and WTP

▶ Willingness to Pay for Private Environmental Goods

## Real Investment Evaluation

▶ Project Evaluation

## Reasoning and Economic Rationality

▶ Economic Rationality Assumption

## Recalibration Response Shift

▶ Response Shift

## Receptive Awareness

▶ Mindfulness and Life Satisfaction

## Reciprocal Causal Relationships

▶ Bidirectional Explanation of Satisfaction

# Reciprocal Causation

Melanie Revilla
Political and Social Sciences, Universitat Pompeu Fabra RECSM, Barcelona, Spain

## Synonyms

Reciprocal interaction

## Definition

Reciprocal causation occurs when two or more events have simultaneously causal effects upon each other.

## Description

Causation (also referred to as causality) is the relationship between two events (variables), where the second one, called effect, is a consequence of the first one, called cause.

Causation has to be distinguished from correlation: a correlation between two variables only means that the variables are related but does not indicate the nature of the relationship.

In the case of causation, if the first event does not occur, then the second event cannot occur either. The occurrence of the first event is a necessary condition for the occurrence of the second. Usually, the cause, or independent variable, is denoted X, whereas the effect or dependent variable (it depends on the occurrence of X) is denoted Y.

Most often in social sciences research, it is assumed that the relationship between Y and X is unidirectional and that the researcher knows the direction of the effect: we know that it is X that influences Y, so X is the cause and Y the effect, not the contrary. For some variables, this assumption is not at all problematic. For instance, if the researcher is interested in the relationship between gender and life satisfaction, there is no doubt that if there is causation, gender has to be the cause and life satisfaction the effect. The contrary does not make sense. The same apply if the researcher is looking at the relationship between age and political participation, between race and perceived discrimination, etc. The direction of the causation is also clear in time series analyses: a variable measured at time $t$ cannot affect a variable measured at time $t-1$. The temporal ordering between variables is clear: the potential causation's direction too.

However, there are also cases where the direction of the causation is not so clear and cases where it seems theoretically justified assuming that not only X affects Y but at the same time Y also affects X.

*Reciprocal causation* refers to this exact situation where two events influence each other simultaneously. X is both a cause of Y and an effect of Y. It is both an independent variable that explains Y and a dependent variable that is explained by Y. The same is true for Y. The reciprocal causation is also called bidirectional causation.

Reciprocal causation occurs in all kinds of scenarios, some of them being part of our everyday life.

One often given example is the one of a car accident where two cars collide frontally because of the reactions of each conductor to the action of the other conductor. Two cars are driving in opposite directions. One driver starts to drive in the wrong lane. When seeing that, the other driver, to avoid a crash, reacts by changing his direction and starts driving in the other car's lane. But simultaneously, the first driver realizes that he was driving in the wrong lane and that a car is coming in front of him, so he starts moving back to his own lane. And now the second driver when seeing the first driver is going back to his lane wants to come back to his own lane. But simultaneously the first driver that saw the second driver switching direction switches too. This continues till they both ultimately crash. The accident is the result of one mistake of the first driver when first changing lane and of a series of simultaneous defensive driving actions where both drivers' decisions are affected by what the other is doing.

Reciprocal causation models are widely used and applied to behavior and social learning theories. In the field of behavior analyses, it is indeed often assumed that behaviors of different

agents co-occur such that each agent's behavior causes the other agents' behavior and at the same time is affected by the other agents' behavior. In the field of social learning, Albert Bandura (1986) proposes a theory of reciprocal determinism. He uses the term of "triadic reciprocal causation" to refer to the mutual influence between three variables: environment, behavior, and person (mainly cognitive factors as memory or anticipation). Bandura underlines that these three variables can affect each other. It is the interaction between these three variables that result in human actions.

More generally, reciprocal causation models are used in all different kinds of disciplines, from physics to economics, going through sociology. We will only give one last example: Quinn and Duckworth (2007) studied the relationship between well-being and academic performance. They ask themselves: "Does school achievement come at the expense of happiness? Or, conversely, are better-performing students happier?" The direction of the causation between the two variables is not clear. In fact, they conclude that the relationship between well-being and academic performance may well be a reciprocal causation.

## References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Quinn, P. D., & Duckworth, A. L. (2007, May). *Happiness and academic achievement: Evidence for reciprocal causality*. Poster session presented at the annual meeting of the American Psychological Society, Washington, DC.

## Reciprocal Interaction

▶ Reciprocal Causation

## Reciprocity Among Researchers and Subjects

▶ Research Relationship(s)

## Reciprocity in Exchange

▶ Exchange Theory

## Reclaimers of Waste

▶ Street Waste Pickers in Pretoria, South Africa

## Reconceptualization Response Shift

▶ Response Shift

## Recovering Societies

▶ Post-conflict Societies

## Recovery

▶ Resilience

## Recreation

▶ Play Behavior

## Recreational Child Care

▶ Parental Time and Child Well-Being

## Recurrent Education

▶ Lifelong Learning

## Recyclers of Street Waste

▶ Street Waste Pickers in Pretoria, South Africa

## Redistribution of Economic Resources

▶ Income Redistribution

## Redistribution of Income

▶ Income Redistribution

## Redistributive Policies

▶ Income Redistribution

## Reduced Energy Levels

▶ Fatigue

## Reduced Muscle Strength

▶ Fatigue

## Reduced Work Hours

▶ Part-Time Work
▶ Workplace Flexibility

## Redundancy and Work

▶ Unemployment

## Reengineering the Public Sector

▶ Service Quality in New Public Management

## Reference Drift

▶ Preference Drift

## Reference Guidelines

▶ Normative Data

## Reference Period

Diana Zavala-Rojas
Research and Expertise Centre for Survey Methodology, Department of Political and Social Sciences, Universitat Pompeu Fabra, Barcelona, Spain

### Synonyms

Time reference

### Definition

The time reference in a survey item. Requests can be asked about a present situation, future events, and intended behavior. Survey items can also be directed to past events or actions.

### Description

Information requested in a survey is often bounded into a time point, although some questions might be directed to characteristics that never change or are permanent such as place of birth, name, and

native language. When survey items are asked about the past, the time indicated in the request is called *reference period*. These items are largely used in ▶ survey research. Many questions in survey research ask about the past or what is called *autobiographical memory*. Schwarz and Sudman (1996) distinguish three types of autobiographical memory. One is about *personal memories* (a mental image of a particular event, feeling, or thing). The second type is *autobiographical fact* that implies remembering a fact but without having a mental image of it. And the third type is *generic personal memory* that means having a mental image of what an event is, but without any image of a specific point in time. They said that people store information in hierarchical order and that some features in the survey design can provide guide to retrieval.

However, questions about past information are also more complicate and can be a burden for respondents Tourangeau et al. (2000). Saris and Gallhofer (2007) state that items asking about the past can have different problems depending on how the reference period is formulated. Respondents are less likely to produce accurate responses if they are asked about very long reference periods. Respondents find also problematic to count events in the past, especially if they are asked about activities that occur very frequently, for example, time spent watching television. They also found that questions about the future have a negative systematic effect in the ▶ reliability and also in the validity. Schuman and Presser (1981) showed that people have a tendency to see events closer to the date of the interview as they were in reality. Scherpenzeel (1995) found that people estimated better the occurrence of events in a point or period of time if they are asked in a two-step procedure, using first a larger reference period and then the reference period of interest of the researcher.

## Cross-References

▶ Method Effects
▶ Reliability
▶ Respondent Burden
▶ Survey Research

## References

Saris, W. E., & Gallhofer, N. I. (2007). *Design, evaluation and analysis of questionnaires for survey research*. Hoboken, NJ: Wiley-Interscience.
Scherpenzeel, A. (1995). *A question of quality. evaluating survey questions by multitrait-multimethod studies*. Leidschendam (the Netherlands): KPN Research.
Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys: Experiments on question form, order and context*. New York: Academic.
Schwarz, N., & Sudman, S. (Eds.). (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.
Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, MA: Cambridge University Press.

# Reference Values

▶ Normative Data

# Refugees, Quality of Life

Anna M. Rulska-Kuthy
Department of Political Science and International Affairs, University of North Georgia, Dahlonega, GA, USA

## Synonyms

Displaced persons; Emigres; Forced migrants

## Definition

A refugee is an individual who has left one's own country due to a well-founded fear for his or her own safety. Oftentimes, the internally displaced people are referred to, incorrectly, as refugees. An internally displaced person (IDP) is any person

who has left their residence by reason of real or imagined danger but has not left the territory of their own country. A refugee, on the other hand, is any person who, "owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion, is outside the country of his nationality, and is unable to return to it" (1951 UN Refugee Convention). While refugees are a matter of concern for those in such situations and from a humanitarian perspective, they potentially represent challenges to both their host countries and their country of origin (Marrus, 1985).

## Description

### Causes

Factors that cause individuals to leave their country as refugees vary greatly (Malkki, 1992). Individuals may seek refuge in other countries as the result of persecution on the basis of religion, ethnicity, ideology, and expressing opinions that are seen as dangerous by those in power in their own countries. Refugees might include minority groups that are being systematically targeted for persecution or political dissidents. Groups of people may flee as the result of forced expulsion, such as Jews from fifteenth-century Spain or Germans from post-World War II Poland, but refugees may also gradually flow out of a country that contains an oppressive regime. Many refugees also flee their country as the result of wars, both interstate and especially civil wars (Zolberg, Suhrke, & Aguayo, 1989).

The rise in ethno-national conflict in the post-Cold War world has greatly contributed to the rise in cross-border migration. Deliberate government policies of ethnic cleansing through forced emigration, murder, or harassment can lead to large-scale outflows of people from a country (e.g., Bosnia, Kosovo, Rwanda). Governments have enacted such policies at times in an attempt to achieve either ethnic homogeneity or dominance, to eliminate political dissidents, or in order to provide scapegoats to the majority population for failures of state policies (Levy, 2001). This can also happen as a result of intercommunal conflict within a country that does not happen as a result of state action (e.g., Iraq in the twenty-first century).

Millions of refugees are currently dislocated from their countries due to these conflicts. They have fled from their home country and its conflict for their own safety either due to the general increase in insecurity or specific fear for their security. Many people flee their homes due to conflict but have not fled their country. However, the great displacement of people across international borders has led numerous governments and scholars to consider the impact of refugees in the framework of national and international security rather than simply as a social and/or humanitarian problem. This flow of refugees has become entangled with conflict both within and between states (Levy, 2001).

Immigrant populations that happen as the result of refugee flows can contribute to conflict both within and between states in a number of ways. Sudden increases in immigrant populations can lead to social conflict in host countries by placing strain on limited resources, competing for jobs with local populations, or appearing as a threat to the social identity of some in the host country. Migrants can contribute to conflict between the host and donor country by serving as a point of focus in relations between the countries (Levy, 2001). Whether as a result of difficulties dealing with social problems arising from refugee flows, humanitarian concerns, or pressure from domestic populations from the country of the refugees' origins, the situation can lead to tension between the host and donor countries, which can potentially lead to conflict. For example, refugee flows from Haiti in the early 1990s played a large role in pushing the United States government to eventually intervene through political, economic, and military means to restore President Aristide, who had been ousted in a coup, to power.

### Effects on Countries of Origin

Refugee flows can be challenging for both host countries and for the countries from which they originated. Large-scale flows of refugees out of a country can drain it of human capital that may

be critical to its economy, especially since it is most likely that individuals who are educated or have significant resources are the most likely ones to have the means to flee. Significant refugee flows can also bring international attention to the problems that caused them, making the government in the country of origin look bad. Especially in cases where refugees are fleeing oppressive regimes, they can work to bring attention to human rights abuses being carried out in their home countries. These flows can create potential political and social problems in host countries, which can lead these host countries to intervene in order to stem the flows of refugees, whether politically, economically, or even militarily, such as the US intervention in Haiti in the early 1990s. Refugee populations can potentially fuel conflicts in their country of origin by funding insurgencies through remittances, such as in the case of Kosovo or Sri Lanka (Pirkkalainen & Abdile, 2009).

### Effects on Host Countries

Host countries can face various challenges in dealing with large influxes of refugees. This can place strains on resources, including social services within the host country. In addition, refugees might compete with low-income populations in the host country for these services and employment. In some cases, large refugee flows can contribute to an increase in violent crime activity, such as was the case with a small number of the refugees from El Salvador's civil war. In some rare cases, refugee flows can contribute to the spread of conflict into these new countries. This can especially be the case when the host country and country of origin border one another and the conflict the refugees are fleeing is fought along ethnic lines that span borders, such as in the case of Africa's Great Lakes Crisis (Prunier, 2009).

### Protections for Refugees

For many years, refugees did in fact exist, driven to leave their countries due to wars or persecution, but there was no comprehensive framework internationally to deal with refugees or provide protections for them (Grahl-Madsen, 1966). Following the experience of World War II and the Holocaust, there was increase focus on the rights of refugees. In 1951, the United Nations Convention on the Status of Refugees was written, eventually being signed by 147 countries. According to the UN Convention, signatory countries must:

1. Not return someone to a country where she/he would be in danger.
2. Offer protection to refugees by granting the same rights as other foreigner.
3. Not punish refugees who have entered the country illegally (1951 UN Refugee Convention).

In addition to the general aspirational goals of the treaties, it set certain basic rights of refugees and basic standards for their treatment. If the threat that caused refugees to flee has abated (such as the end of the war), there is the issue of refugee repatriation. The treaty also charged the United Nations with the responsibility to help people who were persecuted in their home country or who feared returning to it (Rourke, 2008). Efforts to deal with issues relating to refugees through the United Nations framework eventually led to the creation of the United Nations High Commission on Refugees (UNHCR). The mission of this agency is to oversee the provision of food and shelter to refugees and to work toward the eventual goal of restoring them to their homes should the situation the refugees fled improve. The UNHCR is one of the largest organizations within the United Nations (Rourke, 2008).

Asylum is legal protection that is given to refugees that protects them from being returned to their home country against their will where they might face persecution at the hands of government agents or other groups. Furthermore, asylum offers protection and immunity from extradition granted by a government to a political refugee from another country (1951 UN Refugee Convention).

Recognition of a refugee's right to asylum is determined entirely by his or her host country through its own legal process for such matters. In some cases, refugees may not have access to these processes where they are not well established or strongly functioning or because of large and sudden refugee flows with which the system cannot deal. Refugees may also

choose to forgo these proceedings because obtaining asylum is far from guaranteed, even in many legitimate cases, in part due to the politically sensitive nature of a host country legally recognizing individuals from another country as refugees that have genuine reasons to fear for their safety.

### Quality of Life for Refugees

In the immediate period after fleeing dangerous situations, refugees often end up in transitional locations while awaiting permanent resettlement. In these refugee camps, which sometimes hold populations in the tens of thousands, people face many great challenges. This includes as high levels of disease and lack of permanence for residents that sometimes live there long term (De Vries & Van Heck, 1994). In these unstable situations, refugees can languish in these conditions for years waiting for resettlement in a third country or for the dangerous conditions in their home country to abate. While these refugee camps are meant to be temporary places for people to reside, they sometimes continue for years and even decades. Some camps that were set up for Palestinian refugees have been operating for over half a century (CBC News, 2007). Even if refugees are able to make it to more permanent dwellings in third countries, they can find difficulty with adjusting to what is often a completely foreign culture, language, and society. This can cause great difficulty in finding employment and being able to access basic services and navigate new societies. Without assistance, whether from government agencies, nonprofit organizations, or communities in the host country, refugees can have great difficulties with adjusting to the new society, including difficulty finding gainful employment (Orley & Kuyken, 1994).

### Refugee Repatriation

In cases where the circumstances that caused refugees to leave their country have abated, refugees often return, whether through their own initiative or because in such circumstances their refugee status can potentially be revoked. While the right of refugees to return to their homeland following conflict is quite reasonable, it can create great strain on the post-conflict societies to which they are returning. A certain degree of refugee repatriation is natural following the end of the conflict that caused people to flow. Its relationship with peace agreements varies, however. Some see it as simply something that will be allowed and will naturally occur to some degree. Others see potential large-scale refugee repatriation as problematic to stability in the post-conflict. However, eventual repatriation of at least a large portion of the refugee population is often perceived as evidence of the success of peace accords. Others sometimes take the harder stance that refugee repatriation is a necessary condition and part of the peace settlements in the case of civil wars or other extended conflicts. However, this repatriation and resettlement can pose a separate and new challenge to the peacemaking and post-settlement peace, especially in cases where the former lands or properties of refugees has been occupied by others in the interim (Adelman, 2002).

## Cross-References


- ▶ Human Rights
- ▶ Immigrants, an Overview
- ▶ Immigrants, Responses to
- ▶ Migration, an Overview


## References


Adelman, H. (2002). Refugee repatriation. In S. J. Stedman, D. Rothchild, & E. M. Cousens (Eds.), *Ending civil wars: The implementation of peace agreements*. Boulder, CO: Lynne Rienner Publishers.

Anatomy of a refugee camp. Canadian Broadcast (CBC) News. June 19, 2007. Available at: http://www.cbc.ca/news/background/refugeecamp/

Convention and Protocol Relating to the Status of Refugees. Text of the 1951 Convention Relating to the Status of Refugees. Text of the 1967 Protocol Relating to the Status of Refugees. Resolution 2198(XXI) adopted by the United Nations General Assembly. Available at: www.unhcr.org/3b66c2aa10.html

De Vries, J., & Van Heck, G. L. (1994). Quality of life and refugees. *International Journal of Mental Health, 23*(3), 57–75.

Grahl-Madsen, A. (1966). *The status of refugees in international law*. Leyden, Germany: A.W. Sijthoff.

Levy, J. S. (2001). Theories of interstate and intrastate war: A levels-of-analysis approach. In C. A. Crocker, F. O. Hampson, & P. Aall (Eds.), *Turbulent peace: The challenges of managing international conflict*. Washington, DC: United States Institute of Peace Press.
Malkki, L. (1992). National Geographic: The rooting of peoples and the territorialization of national identity among scholars and refugees. *Cultural Anthropology, 7*(1), 24–44.
Marrus, M. (1985). *The unwanted: European refugees in the twentieth century*. New York: Oxford University Press.
Orley, J. H., & Kuyken, W. (Eds.). (1994). *Quality of life assessment: International perspectives: Proceedings of the joint meeting organized by the world health organization and the foundation IPSEN in Paris, July 2–3, 1993*. Berlin: Springer.
Pirkkalainen, P. & Abdile, M. (2009). The diaspora-conflict-peace-nexus: A literature review. Diaspeace Working Paper No. 1.
Prunier, G. (2009). *Africa's world war: Congo, the Rwandan genocide, and the making of a continental catastrophe*. New York: Oxford University Press.
Rourke, J. (2008). *International politics on the world stage* (12th ed.). New York: McGraw Hill.
Zolberg, A. R., Suhrke, A., & Aguayo, S. (1989). *Escape from violence: Conflict and the refugee crisis in the developing world*. New York: Oxford University Press.

## Regeneration of Neighborhoods

► Gentrification

## Regenerative Cities

► Sustainable Communities Movement

## Regime Thesis About Welfare State Attitudes

► Welfare State Attitudes

## Regime Types in Asia

► Democracy and Quality of Life in Asian Societies

## Regional Analysis

Bienvenido Ortega
Departamento de Economía Aplicada (Estructura Económica), Universidad de Málaga, Málaga, Spain

### Definition

When an analyst employs spatially disaggregated data (data for units of area) and certain aspects of homogeneity or cohesion, from the perspective of the dimensions of well-being under study, are attributed to these spatial units, these areas are being converted into regions. As a result, the analyst is engaging, by definition, in regional analysis (Isard, 1956).

### Description

Well-being has a regional dimension. This is not only due to the spatial nature of data that may be used in empirical research: regions are defined by particular environmental and socioeconomic factors which are also influenced by the spatial dimension of culture and ► institutions (Odum & Moore, 1938; Plaut, Markus, & Lachman, 2002). In applied research, this means that once the dimensions of well-being have been identified, analysts have to choose the suitable spatial unit of observation and, where appropriate, draw policy-consistent or analytically meaningful territory boundaries.

However, the question of how to define "regions" has caused considerable debate. Contemporary discussions have questioned the traditional and long-established notion of the "region" as a closed, bounded, and territorial entity (Hudson, 2007), arguing that the region is not innate and pre-given in any specific geographical context (Lagendijk, 2007; Pike, 2007). Along these lines, (Eurostat, 2007), for example, distinguishes between normative and functional regions. The limits of normative regions are the expression of political will while functional

regions are defined according to analytical requirements. In this case, areas are grouped together using geographical or socioeconomic criteria. Above all, regions should be defined, from the local to supranational level, so as to be useful for both evaluating the various dimensions of human well-being and ▶ policy analysis.

The criteria to delimit spatial units of analysis might be provided by regionalization methods, whose aim is to aggregate areas into analytical units in such a way that the resulting regions are conveniently related to the phenomena under examination (Duque, Ramos, & Suriñach, 2007). Additionally, in order to decide on the appropriate scale of ▶ spatial analysis, the researcher has to take into account the possible existence of two effects: spatial dependence between variables corresponding to contiguous areas and instability across space in the relationships among variables used to model selected quality of life dimensions, that is, the spatial heterogeneity problem. Spatial dependence exists when there are systematic relationships between observations corresponding to regions that are contiguous in space. This could be due to the existence of a mismatch between the suitable regional scale of the variable under study and the spatial scale at which it is referenced, given the data available. Both exploratory spatial data analysis (ESDA) and maps of local indicators of spatial association (LISA) are instruments that might provide indications of the extent and pattern of spatial association across the data (Anselin, 1995, 1996).

Thus, regional analysis is a key element for both analysts and policy-makers. It is also important in the case of spatial planning, the main focus of which must be to promote direct improvements in the population's ▶ quality of life, a key force in the localization decisions of both individuals and firms (Rappaport, 2009), as indicated by empirical work (Roback, 1982; Rosen, 1979). This was suggested in (Hoover, 1963), which pointed out that an attitude present in every regional analyst is "to appraise the merits of different regions as places to live, places to know, or places to work." This knowledge has to be employed in trying to improve quality of life conditions in a particular area and to design policies aiming at the reduction of existing ▶ regional disparities in well-being.

## Cross-References

- ▶ Institutions
- ▶ Policy Analysis
- ▶ Quality of Life
- ▶ Regional Disparities
- ▶ Spatial Analysis
- ▶ Spatial Planning (Europe)

## References

Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical Analysis, 27*, 93–115.

Anselin, L. (1996). The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In M. Fischer, H. Scholten, & D. Unwin (Eds.), *Spatial analytical perspectives on GIS* (pp. 111–125). London: Taylor & Francis.

Duque, J. C., Ramos, R., & Suriñach, J. (2007). Supervised regionalization methods: A survey. *International Regional Science Review, 30*, 195–220.

Eurostat. (2007). *Regions in the European Union. Nomenclature of territorial units for statistics. NUTS 2006/EU-27*. Luxembourg: Office for Official Publications of the European Communities.

Hoover, E. M. (1963). Whence regional scientists? *Papers and Proceedings of the Regional Science Association, 11*, 7–13.

Hudson, R. (2007). Regions and regional uneven development forever? Some reflective comments upon theory and practice. *Regional Studies, 41*, 1149–1160.

Isard, W. (1956). Regional science, the concept of region, and regional structure. *Papers and Proceedings of the Regional Science Association, 2*, 13–26.

Lagendijk, A. (2007). The accident of the region. A strategic relational perspective on the construction of the region's significance. *Regional Studies, 41*, 1193–1207.

Odum, H. W., & Moore, H. E. (1938). *American regionalism: A cultural-historical approach to national integration*. New York: Holt.

Pike, A. (2007). Editorial: Whither Regional Studies? *Regional Studies, 41*, 1143–1148.

Plaut, V. C., Markus, H. R., & Lachman, M. E. (2002). Place matters: Consensual features and regional variation in American well-being and self. *Journal of Personality and Social Psychology, 83*, 160–184.

Rappaport, J. (2009). The increasing importance of quality-of-life. *Journal of Economic Geography, 9*, 779–804.
Roback, J. (1982). Wages, rents, and the quality-of-life. *Journal of Political Economy, 90*, 1257–1278.
Rosen, S. (1979). Wage-based indexes of urban quality-of-life. In P. Mieskowski & M. Straszheim (Eds.), *Current issues in urban economics* (pp. 74–104). Baltimore: Johns Hopkins University Press.

## Regional Disparities

Bienvenido Ortega
Departamento de Economía Aplicada (Estructura Económica), Universidad de Málaga,
Málaga, Spain

### Synonyms

Regional inequalities

### Definition

Regional disparities in well-being are a component of overall ▶ inequality in quality of life among individuals. These disparities emerge when there are regions that perform better than others in terms of ▶ human development. This then leads to an uneven distribution of quality of life across spatial units. Thus, regional inequalities are present when researchers, through ▶ regional analysis, observe differences in aggregate economic and social indicators of well-being across units or when controlling for personal characteristics, disaggregated data indicates that regional differences in well-being among individuals still remain.

### Description

The regional dimension is relevant in the analysis of well-being. This becomes clear when analysts recognize that, for example, ▶ unemployment reduces well-being in individuals through its spatial effects in quality of life to a greater extent than the simple loss of income effect and touches even those who are not out of work. Thus, regional disparities are a dimension of overall inequality in quality of life but entail added significance when regional divisions align with political, linguistic, religious, or ethnic tensions to undermine social and political stability. In this case, regional disparities may be a contributory factor in social breakdown (Kanbur & Venables, 2005; Murshed & Gates, 2005). Moreover, by analyzing how these differences evolve over time, researchers can study whether regional quality of life is converging (becoming more equal) or diverging (becoming less equal). However, it is important to note that while regional averages may converge, inequalities within regions may be increasing, and, even if regional averages are brought within a common range, nonetheless, total inequality in well-being among the population as a whole could be increasing (Alonso, 1968).

Total inequality in quality of life among individuals in a country where regions have already been defined may be broken down into differences between different regions and existing disparities in well-being within a single region. This can be done when the analyst employs individual (disaggregate) data for people living in those predetermined areas. However, if the ultimate objective of the research is to study interregional differences in the dimensions of quality of life, it is possible to employ aggregate data to characterize regional disparities of different dimensions of well-being in that region, assuming that all individuals in a region have a per capita value (perhaps weighted by the population share). In both circumstances, there are different methods for measuring regional disparities.

One way to tackle regional disparities in well-being is to start from interpersonal disparities and consider their spatial dimensions by using disaggregate data (Kanbur & Venables, 2005). The main aim of this approach is to decompose the national inequality present in each particular quality of life variable into between-regions inequality and within-region disparities. This can be done by using

decomposable generalized entropy class on inequality measures as developed by Shorrocks (1980, 1984). Those measures provide indexes that are additively decomposable into both a within-region and between-regions component. Although such measures have been very widely applied in the analysis of ▶ income distribution among different members of a community, this methodology may also be applied to the analysis of interpersonal inequalities in variables related to human well-being in an aim to disentangle possible sources of regional inequality in quality of life. In addition, this methodology allows us to contrast whether interregional disparities depend on a certain set of determinants different from those explaining intra-regional disparities. Furthermore, this methodology would be useful for policy design, given that when regional disparities dominate total inequality, policy measures can targeted these particular areas. In contrast, dominance of the within-region component calls for policies targeting finer units within these areas, such as distinguishing between rural and ▶ urban areas within a region (Kanbur, Venables, & Wan, 2005).

Notwithstanding, researchers may characterize interregional disparities in quality of life by using aggregate data and assuming that all individuals in a region have the same average or per capita level for each well-being variable considered in the analysis. With this aim, regional analysts may employ raw data and/or achievement and improvement indices constructed for a range of social and economic quality of life indicators. These indices of well-being may be created with the objective of measuring a region's average achievements in relation to the basic dimensions of human development (Anand & Sen, 1994; Kakwani, 1993; OECD, 2008). There are also different statistical tools for measuring spatial inequality, such as sample variance analysis, Pearson's coefficient of variation, or the Gini index. In this context, for example, the ▶ Gini coefficient measures the size of disparities in interregional distribution of a quality of life variable within a country, permitting researchers to point out the dimensions of well-being where regional disparities are more evident and may be detrimental to human development.

Besides this, the study of how regional disparities have evolved over time is a subject in quality of life research. Chenery pointed out that reducing quality of life-related regional disparities is a major concern for regional development policies (Chenery, 1962). In this context, the process by which a specific variable of regional disparity declines over time as regions become more equal is known in the literature as regional convergence, while there is a regional divergence process as well. This is when regions become less equal with respect to the variable being analyzed.

A number of methods are used in the empirical literature to test for the presence of regional convergence in variables. In the field of quality of life research, when analysts are using achievement indices to characterize regional well-being levels, a regional convergence (or catching up) process can be defined when improvements in these ▶ indices of well-being in lagging regions outperform improvements in those that are more advanced. Likewise, divergence occurs when improvements in well-being in more advanced regions outperform improvements in regions that are lagging behind. Furthermore, based on the concept of economic convergence developed in neoclassical growth theory (Solow, 1956), the study of the convergence of variables and/or indicators related to regional well-being may be carried out by using different methods previously employed in empirics of economic growth: analyzing the correlation between the initial levels and subsequent growth rates of the variable under study (a concept known in the literature as β-convergence) and the later study of whether there is a reduction over time in the dispersion of the variable being (σ-convergence).

Specifically, regional β-convergence exists in a variable when this variable tends to grow in lagging regions at a faster rate than in more advanced ones. We can test for this, according to Baumol (1986), by using regression methods to estimate (Eq. 1):

$$\ln Y_{it} - \ln Y_{it-1} = a + b \ln Y_{it-1} \qquad (1)$$

where i (=1, . . ., N) refers to the regional unit, t (= 1, . . ., T) to the year, and Y is a (per capita) social, economic, or quality of life indicator. Parameter (a) is a constant term across regions, and parameter (b) captures the effect of each indicator's initial level on its rate of growth as

$$b = 1 + \exp(-\beta) \quad (2)$$

where β represents the speed of convergence to a common level of equilibrium. Consequently, rejecting the null hypothesis that the ▶ regression coefficient $b = 0$ against $b < 0$ in (Eq. 1) implies β-convergence (i.e., that relatively lagging regions have improved the most and are catching up with the leading regions). Meanwhile, the concept of σ-convergence refers to the evolution of a variable's cross-sectional dispersion over time. From this perspective, convergence exists if the ▶ dispersion of the indicator being studied (measured, e.g., by the standard deviation of the log of the indicator or by the coefficient of variation) diminishes over time (Barro & Sala-i-Martin, 1999). In addition, rate of σ-convergence may be calculated as the annual percentage of change between a measure of dispersion at a given moment in time (t + T), and this measure calculated at moment (t), where a negative (positive) value implies σ-convergence (divergence).

However, a variety of additional analytical methods have been proposed in order to consider the dynamics of the whole distribution of the variable (across regions and over time), not merely its average behavior captured through β- and σ-convergence methods. Along this line, Bianchi (1997), Quah (1993) employ kernel density estimates for different years in order to assess the likelihood of convergence of a continuous distribution. In this context the stochastic kernel shows how likely it is for one region to move from one specific level at point (t) in time to a certain level at point (t + 1), for example, whether regions that perform below average in (t) are expected to perform better at period (t + 1). This methodology makes it possible to point out the presence of different convergence dynamics such as polarization and clustering (Hobijn & Franses, 2001; Neumayer, 2003).

There are real geographical features that can affect regional disparity in quality of life (Escobal & Torero, 2005; Gallup, Sachs, & Mellinger, 1999). Notwithstanding, the literature shows that there are also additional determinants of such disparities, particularly in developing countries, such as quantity and quality of regional infrastructure endowment or quality of regional institutions and other political factors (Easterly, Ritzan, & Woolcock, 2006). Moreover, as previously mentioned, the traditional focus on interregional disparities has been complemented in the quality of life literature with intra-regional analysis, in which the study of inequalities using the urban-rural dimension is of particular importance, especially in developing countries. Along this line, (Sahn & Stifel, 2003) conducted tests for urban-rural convergence in Africa regarding achievement indices for eight different welfare indicators. They conclude that more remote areas benefit less from growth in terms of impact on well-being. (Escobal & Torero, 2005) emphasize that urban areas in Peru tend to have agglomeration forces act to generate virtuous circles of self-reinforcing development. (Gräb & Grimm, 2011) also conclude that communities in Burkina Faso are poor, not only because the households living there have characteristics which make them poor but also because these communities are geographically disadvantaged, provide fewer public services and offer only limited access to markets and infrastructure.

In sum, the capacity to explain where regional disparities in quality of life come from and how they evolve over time is crucial in the design of policies aimed at facilitating lagging regions opportunity to share in benefits of faster national development, in both developed and ▶ developing countries. This is a real challenge for quality of life researchers at the regional level.

## Cross-References

- ▶ Developing Countries
- ▶ Dispersion
- ▶ Gini Coefficient
- ▶ Human Development
- ▶ Income Distribution

- ▶ Indices
- ▶ Inequality in Quality of Life
- ▶ Regional Analysis
- ▶ Regression Coefficient
- ▶ Unemployment
- ▶ Urban Areas

## References


Alonso, W. (1968). Urban and regional imbalances in economic development. *Economic Development and Cultural Change, 17*(1), 1–14.

Anand, S., & Sen, A. (1994). *Human development index: Methodology and measure*. New York: Human Development Report Office. United Nations Development Programme, Occasional Paper 12.

Barro, R. J., & Sala-i-Martin, X. (1999). *Economic growth*. Boston: The MIT Press.

Baumol, W. J. (1986). Productivity growth, convergence and welfare: What long-run data show. *American Economic Review, 76*, 1072–1085.

Bianchi, M. (1997). Testing for convergence: Evidence from non-parametric multimodality tests. *Journal of Applied Economics, 12*, 393–409.

Chenery, H. B. (1962). Development policies for Southern Italy. *Quarterly Journal of Economics, 76*, 515–547.

Easterly, W., Ritzan, J., & Woolcock, M. (2006). Social cohesion, institutions, and growth. *Economics and Politics, 18*(2), 103–120.

Escobal, J., & Torero, M. (2005). Adverse geography and differences in welfare in Peru. In R. Kanbur & A. J. Venables (Eds.), *Spatial inequality and development* (pp. 77–122). Oxford, UK: Oxford University Press.

Gallup, J. L., Sachs, J., & Mellinger, A. (1999). Geography and economic development. *International Regional Science Review, 22*(2), 179–232.

Gräb, J., & Grimm, M. (2011). Inequality in Burkina Faso - to what extent do household, community and regional factors matter? *Journal of the Royal Statistical Society, 174*(3), 1–26.

Hobijn, B., & Franses, P. H. (2001). Are living standards converging? *Structural Change and Economic Dynamics, 12*, 171–200.

Kakwani, N. (1993). Performance in living standards. An international comparison. *Journal of Development Economics, 41*, 307–336.

Kanbur, R., & Venables, A. J. (2005). *Rising spatial disparities and development*. Policy Brief, number 3. United Nations University. http://archive.unu.edu/publications/briefs/policy-briefs/2005/spatial-inequality.pdf. Accessed 15 March 2013.

Kanbur, R., Venables, A. J., & Wan, G. (2005). Spatial disparities in human development: An overview of the Asian evidence. In R. Kanbur, A. J. Venables, & G. Wan (Eds.), *Spatial disparities in human development: Perspectives from Asia* (pp. 1–6). Tokyo: United Nations University Press.

Murshed, S. M., & Gates, S. (2005). Spatial horizontal inequality and the Maoist insurgency in Nepal. In R. Kanbur, A. J. Venables, & G. Wan (Eds.), *Spatial disparities in human development: Perspectives from Asia* (pp. 188–206). Tokyo: United Nations University Press.

Neumayer, E. (2003). Beyond income: Convergence in living standards, big time. *Structural Change and Economic Dynamics, 14*, 275–296.

OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD Publishing.

Quah, D. (1993). Galton's fallacy and tests for convergence hypothesis. *The Scandinavian Journal of Economics, 95*(4), 427–443.

Sahn, D., & Stifel, D. (2003). Urban-rural inequality in living standards in Africa. *Journal of African Economies, 12*(1), 564–597.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica, 48*, 613–625.

Shorrocks, A. F. (1984). Inequality decomposition by population subgroup. *Econometrica, 52*, 1369–1385.

Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics, 70*, 65–94.


## Regional Hypotheses

- ▶ Faceted Action System Theory (FAST)
- ▶ Faceted Smallest Space Analysis (Faceted SSA; FSSA)

## Regional Inequalities

- ▶ Regional Disparities

## Regional Mobility

- ▶ Internal Migration

## Regional Planning

- ▶ Land-Use Planning
- ▶ Planning, an Overview
- ▶ Planning, Spatial

# Regional Quality of Life

Annette Spellerberg[1], Roland Habich[2] and Denis Huschka[3]
[1]University Kaiserslautern, Kaiserslautern, Germany
[2]Data Management, Social Science Research Center (WZB), Berlin, Germany
[3]Rat für Sozial- und Wirtschaftsdaten, Berlin, Germany

## Definition

Regional ▶ quality of life reflects levels of regional disparities within different countries, marking privileged and handicapped regions with respect to standards and living and individual well-being.

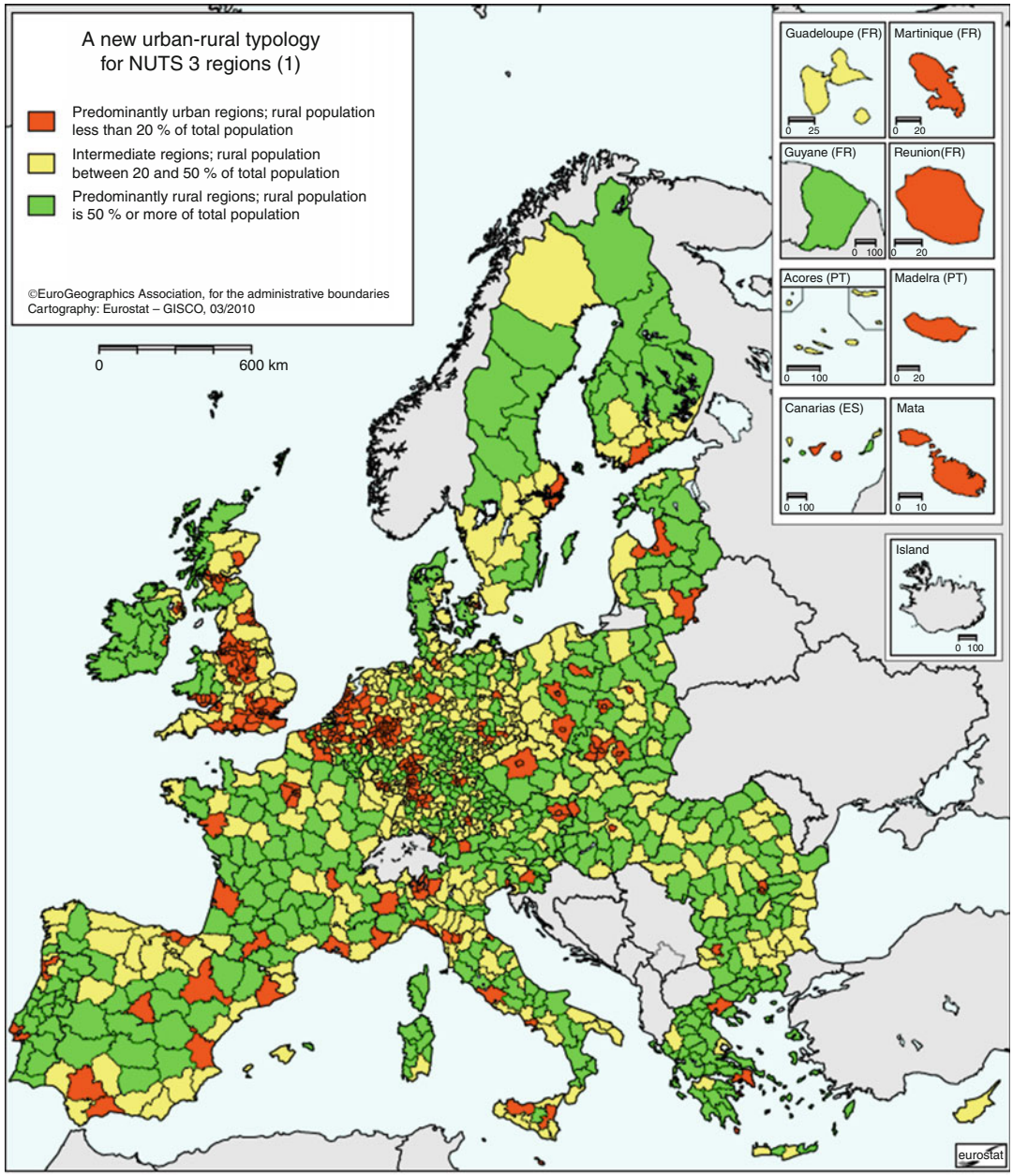
## Description

Regional phenomena are situated at a middle level between the local community and the nation as a whole. From an individual point of view, regions can be understood as a constructed space, as an interlinked space in economic, geographic, and cultural rather than in administrative terms. Because there are not any databases dealing with boundaries of regions as perceived individually, we focus on regions on the basis of administrative boundaries. In view of the socioeconomic structural changes that are developing towards a transnational world, an increasing importance has come to be attributed to the regional scale. State regulations or EU transfers follow the guideline to counterbalance regional disparities, but they succeed only to a certain degree. The regional level has become more important with respect to growing disparities, standard of living, and social inequality (Heidenreich & Wunder, 2008).

Successful regions are seen as drivers of economic development, as good examples of networking and ▶ social integration, and as innovation centers for the development of new technology (Rhodes, 1995). In peripheral regions, we find economic weakness as well as outmigration, a strongly ageing population, devaluation of dwellings, and loss of infrastructures and services such as schools, libraries, hospitals, medicines. As far as technical infrastructure is concerned, broadband Internet is a key issue for economic recovery, but is not available in sparsely populated and therefore unprofitable places. On the other hand, regional communities can act as a counterweight to the homogenization and dislocation brought about by international economic processes (Giddens, 1986). Regional culture and traditions may help to strengthen a sense of social identity, which constitutes a contrast and an anchor in relation to the influence of international media, global lifestyles, and uncertainties about future social developments. These regional developments indicate different aspects of quality of life, standard of living, and potentials of development as well as social ties and ▶ sense of belonging.

The consistent yardstick to distinguish regions is the population factor, rural regions being related to low levels of settlement, whereas central or metropolitan areas are marked by high population density. Within the EU-27, four out of ten inhabitants live in urban areas and 36 % in intermediate areas. Only 24 % of population lives in rural areas, whereas the rural share of land area encompasses 56 %. However, the share of population in predominantly rural or urban region differs significantly within countries. Ireland hosts the lowest, and Malta, Cyprus, the UK, and the Netherlands host the highest share of people living in cities (Eurostat, 2010; Fig. 1). In these different types of settlements, we find that different conditions of living, but empirically based information of regional disparities, in terms of subjective perceptions and well-being are hardly available.

Because of the lack of data, empirical research on regional quality of life is less often conducted than on the community, national, or international level. While statistical data on living conditions are collected for defined administrative units at local, regional, and national level, perceptions and ▶ subjective well-being are not included in official data bases. Representative national and internationally administered surveys like




(1) This typology is based on a definition of urban and rural 1km$^2$ grid cells. Urban grid cells fulfil two conditions: 1) a population density of at least 300 inhabitants per km$^2$ and 2) a minimum population of 5000 inhabitants in contiguous cells above the density threshold. The other cells are considered rural. Thresholds for the typology: 50% and 20% of the regional population in rural grid cells.

For Madeira, Açores and the French outermost regions, the population grid is not available. As a result, this typology uses the OECD classification for these regions.

**Regional Quality of Life, Fig. 1** The EU urban-rural typology for NUTS 3 regions (Source: Eurostat (2010, p. 241))

ISSP, European Quality of Life Survey, Gallup World, and East Asian Social Survey inform about ► subjective indicators, but usually these datasets do not reflect regional sampling (compare: The Research Data Center "International Survey Programs"; http://www.gesis.org/en/institute/competence-centers/rdc-international-survey-programmes).

In most studies on regional disparities data about demographic trends and migration, economic power and infrastructure (especially health issues) cast light on the discrepancies that separate different regions. The lowest and highest employment rates differ about 40 % points in 2008 in the EU-27, the regional unemployment rates differed accordingly; in the south of Spain or in the north of East Germany, unemployment rates are above 15 % (Eurostat, 2010, p. 51) and employed persons with higher education are found most of all in Inner London (55 %) and least of all in rural Czech Republic and Romania (8 %). Access to broadband Internet shows huge differences with the highest share in the north of the EU (Sweden, the UK, and the Netherlands) and lowest share in Bulgaria, Romania, and Greece (84 % vs. 20 %). On the level of the economy and in relation to other objective aspects, the existing regional differences between the different EU member states and different regions are considerable. We face an increasing divergence between the central urbanized and peripheral rural regions in spite of subsidization policies. The same holds true for the states of the USA, Asia, and other parts of the world.

In comparative studies, ▶ life satisfaction is analyzed for different countries; deeper analysis (NUTS 1-NUTS 3 regions) can hardly be conducted in comparable datasets (Watson et al., 2010). In multivariate analysis, social scientists compare the influence of demographic and socio-economic factors, personal health, and perception of the quality of society on subjective well-being. Regional variations normally can be included only by means of the variable "size of community" as an indicator for the urban-rural continuum.

Subjective indicators enable us to check whether differences in the standard of living as found in various regions are associated with different levels of individual well-being. Compensation effects and comparison processes affect subjective well-being, which results in the fact that satisfaction levels in rural regions are not notably worse than in urbanized areas – despite differences in objective living conditions. Normally, we find corresponding results: people that are better off are more satisfied with living conditions than people with lower living standards. Comparing urban and rural regions within the EU, paradox results of low level of living standards and high satisfaction can be found. Comparing rural and urban regions in Germany, data show that the inhabitants of small villages or farming communities (as a proxy) more often evaluate their financial position more positively than persons living in major cities and their suburbs, because costs of living are lower, especially costs for housing (e.g., in Austria, Belgium, and East Germany). The same results can be shown for ▶ happiness and general satisfaction with life in Austria, the UK, or the Netherlands. In the sparsely populated Scandinavian countries like Sweden or Finland, the perceptions between urban and rural population do not differ much (Spellerberg, Huschka, & Habich, 2007, p. 301).

According to Veenhoven (2002), happiness strongly correlates with occupational prestige, but other factors influencing happiness are participation in associations, having a spouse, having friends, material aspects of life, as well as a sense of ability to control one's environment. Individuals are able to weight the disadvantages of rural living conditions (labor market and education) and advantages of a more compact social world. However, actual data of East Germany show that rural regions fall behind and that the advantages are not any longer able to prevent dissatisfaction with the place of residence due to lack of perspectives (Habich & Spellerberg, 2011). What follows is that better educated and younger people migrate to places that promise better earning opportunities and professional prospects.

## Conclusion

From an economic and demographic point of view, peripheral rural regions are at risk of falling into a descending spiral. Structurally weak areas located in peripheral regions suffer from de-economizing and depopulation as well. Standard of living is significantly lower, regarding income, job opportunities, unemployment, and medical infrastructure. The economic sphere is evaluated less favorable and life satisfaction, satisfaction with family life, housing conditions, and quality of the environment more favorable than in urban regions. The discrepancies

between town and country in terms of subjective well-being are not as marked as might have been expected on the evidence of objective indicators.

The findings imply that there are remarkable discrepancies in terms of life quality in different regions within countries as well as in Europe as a whole. We have to find a new approach to living with increasing social disparity, from which the national level is more important for individual comparisons as yardstick for quality of life. If social indicators are to be able to measure ▶ welfare and monitor ▶ social change and social inequality, we will need regional representative data that is internationally comparable.

## Cross-References

▶ Population Density
▶ Regional Disparities

## References

Eurostat (Ed.). (2010). *Regional yearbook (2010)*. Accessed August 30, 2011, from http://epp.eurostat.ec.europa.eu/portal/page/portal/publications/regional_yearbook

Giddens, A. (1986). *Constitution of society: Outline of theory of structuration*. Berkeley, CA: University of California Press.

Habich, R., & Spellerberg, A. (2011). Regionale Lebensqualität. Appearing in: Statistisches Bundesamt, WZB (Hg.): Datenreport 2011. Wiesbaden: Statistisches Bundesamt.

Heidenreich, M., & Wunder, C. (2008). Patterns of regional inequality in the enlarged Europe. *European Sociological Review, 24*(1), 19–36.

Rhodes, M. (Ed.). (1995). *The regions and the New Europe: Patterns in core and periphery development*. Manchester: Manchester University Press.

Spellerberg, A., Huschka, D., & Habich, R. (2007). Quality of life in rural areas: Processes of divergence and convergence. *Social Indicators Research, 83*, 283–307.

Veenhoven, R. (2002). Why social policy needs subjective indicators. *Social Indicators Research, 58*, 33–45.

Watson, D., Pichler, F., & Wallace, D. (2010). *Second European quality of life survey. Subjective well-being in Europe. European foundation for the improvement of living and working conditions*. Dublin. Accessed August 30, 2011, from http://www.eurofound.europa.eu/pubdocs/2009/108/en/1/EF09108EN.pdf

# Regression

▶ Ordinary Least-Squares (OLS) Model
▶ Random Effects Regression for Panel Data
▶ Univariate Tests

# Regression Coefficient

▶ Beta Weights

# Regression Coefficients

Ralph L. Piedmont
Pastoral Counseling Department, Loyola University Maryland, Columbia, MD, USA

## Synonyms

Beta coefficient; Intercepts; Partial coefficients; Slope; Unstandardized coefficients

## Definition

Regression coefficients are partial coefficients that they indicate the unique predictive contribution of a variable, independent of the influence of every other variable in the regression analysis (Tabachnick & Fidell, 2001).

## Description

The regression analysis is the statistical method that summarizes the direct relationship between the outcome and a set of explanatory variables, calculating for each explanatory variable its "net influence" in explaining and predicting the outcome. When the regression equation is univariate and linear ($Y = a + bx$), the regression coefficients are the intercept (a) that represents the weight long-run value of Y and the slope

of regression and (b) that represents the correlation (rate of change) of the variable (Y) as a function of changes in the other (X) (Cohen & Cohen, 1983).

The values of regression coefficients are dependent on the metric of measurement of the corresponding variables. Thus, a regression coefficient represents the expected amount of change in the dependent variable for a one-unit change on its associated predictor. The regression coefficients are valuable for examining the extent to which particular variables predictive strength can be replicated in the same model across samples.

### Discussion

The estimation of regression coefficients can be done with various methods of estimation under certain statistical assumptions. The most common methods of estimating the unknown regression coefficients are the ordinary least squares (OLS) and the maximum likelihood (ML). Even when a regression coefficient is correctly interpreted as a rate of change of a conditional mean, it is important to take into account the uncertainty in the estimation of the regression coefficient. Therefore, various statistical tests have been developed in order to evaluate the significance of a regression coefficient's value.

## Cross-References

- ▶ Beta Weights
- ▶ Least Squares Regression Line

## References

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon.

# Regulations for Land Use

- ▶ Zoning

## Related EPIC Short Form

- ▶ Expanded Prostate Cancer Index Composite (EPIC)

# Relational Contributions to Optimal Sexual Experiences

Peggy J. Kleinplatz[1], A. Dana Ménard[1] and Meghan Campbell[2]
[1]School of Psychology, University of Ottawa, Ottawa, ON, Canada
[2]Private Practice, Ottawa, ON, Canada

## Synonyms

Eroticism; Intimate relations and sexual experiences; Optimal sexual intimacy; Pleasure, sexual

## Definition

Relational qualities (e.g., intense eroticism, desire to heighten partner's arousal, high levels of empathy) contribute directly or indirectly to optimal sexual experience.

## Description

Optimal sexual experiences have received almost no attention in the research literature, despite evident interest from the general public and the perceived impact of such experiences on the quality of life for many people. Researchers have identified many of the factors that contribute to sexual dysfunctions but have failed to identify the "causes" of "normal" sexual experiences. Research designed to investigate what relational factors contribute to or set the foundation for someone to experience optimal sexual intimacy has been conducted by the Optimal Sexuality Research Team of the University of Ottawa

(Kleinplatz, Ménard, Paquet et al., 2009; Kleinplatz, Ménard, Paradis et al., 2009).

Some sex therapists (e.g., Broder & Goldman, 2004; Metz & McCarthy, 2007, 2012; Schnarch, 1997) have explored the elements that might contribute to optimal sexual experiences. These include relational qualities, such as trust, empathy, communication, and intimacy (Schnarch; Shaw, 2012). Sex therapists have also suggested that sexual experiences may be enhanced by attention, forethought, and preparation (Broder & Goldman, 2004; Castleman, 2004; Kleinplatz, 2006; Zilbergeld, 2004).

Several major categories of relationship qualities facilitate optimal sexual experiences. Some of these elements pervade the relationship, such as goodwill, trust, and empathy. Other elements are particular to or especially prominent *during* sexual relations, including the intent to heighten passion and erotic intimacy in the moment. Some contributors, such as extraordinarily sensitive and skillful communication, are crucial at all times, and their role in facilitating remarkable erotic intimacy cannot be overstated. The ability to trust and feel secure that one would be safe within the context of sexual relations contributes to optimizing these experiences.

Although Hollywood movies would have us believe that wonderful sex "just happens. . . that it is natural and spontaneous," in reality, optimal sexual experiences require couples to engage in behaviors that set the stage for and create this kind of sex. Crucial elements included mutuality, giving and receiving, and meeting one another's needs and reciprocity; willingness to explore, experiment, and learn, discovery, venturing into the unknown, and taking risks; and communication, sharing, and honesty.

Attitudes and beliefs that contribute to optimal sexual experiences include valuing and nurturing the relationship and seeing sex as a way to further intimacy. Extraordinary lovers recognize that it is necessary to put in effort, demonstrate commitment, and bring deliberate intention to the relationship. In order for optimal sexual experiences to be possible, sex needs to be prioritized, planned for, and time set aside.

Optimal sexual experiences were cherished, valued, and carefully cultivated within participants' relationships, both while they were occurring and in general. This means setting the stage for optimal sexual experiences to occur in the relationship. This extends to setting up an emotionally conducive mood prior to sexual encounters and, finally, to setting up a physical environment congruent with the couple's desires. This entails working in unison, prior to and during optimal sexual experiences through experimentation, exploration, taking risks, communication, mutuality, and intentionality. For example, humor, playfulness, and laughter allow partners to enjoy optimal sexual experiences free of social and emotional constraints. Given considerable time and effort, spontaneity *during* sexual encounters, joy, and imagination all help to make these experiences wonderful.

A variety of feelings within the relationship are important. These include love, liking, caring, positive regard, acceptance, and support, as well as respect and consideration. In addition, the feeling states during sex include a sense of intimacy, connection, and warmth; merger, abandon, and giving self over; unself-consciousness, ability to let go and let loose, and freedom; desire for one another, attraction, and chemistry; and passion and intensity.

Developing the kind of relationship in which optimal sexual experiences are possible requires deliberate choices and cultivation. The relationships in which optimal sexual experiences are embedded are characterized by common values, a mutual sense of equality, and good fit. In addition, knowledge of the partner and his or her body/desires/feelings/erotic wishes is very helpful. These qualities are not static; there must be sufficient emotional maturity of the partners, emotional independence in relationship, and self-soothing and depth to allow for change, evolution, growth, fluidity, and maturing of the relationship in the context of a shared relationship history.

Finally, empathy, the willingness to understand one another at deep levels, can be found throughout the relationship. This level of empathy is characterized by sensitivity, perceptiveness, attention,

tuning in to the other, responsiveness, flexibility, openness, and vulnerability. This involves developing exceptional skill at reading body language and being on the same wavelength as the partner(s). This level of empathy requires that both partners be open and vulnerable with one another and treat each other with care, compassion, consideration, and respect.

Many of these overall relational qualities are also relevant *during* optimal sexual experiences. Heightened pleasure, enjoyment, satisfaction, and intense passion are all important contributors to optimal sexual experiences. Eroticism and the intention to deliberately heighten the quality of the experience as it was happening were crucial. Eroticism involves deliberately heightening the partner's arousal, prolonging the experience, and increasing desire *during* sex (Kleinplatz, 2006). Anticipation, contact, and touch are vital. Optimal sexual experiences require a sense of intentionality in the moment.

Feelings of intense connection and being swept up together effectively shut out the outside world. Being swept up together, beyond distractions, enables couples to let go and abandon themselves to the experience and their partners. Carefully planning such an environment then creates the possibility of flexibility and improvisation within an atmosphere of safety, comfort, and trust. Lovers then communicate throughout the experience in order to adapt and adjust as necessary to one another.

Mutuality and the consensual nature of sexual relations seem so obvious as to be afterthoughts but bear mention: Mutuality (e.g., a mutual sense of connection and being swept up together, mutual feelings of pleasure and enjoyment, mutual respect) seemed to be almost woven into the definition of optimal sexual experiences. The need for mutual consent is a basic and fundamental prerequisite for optimal sexual experiences.

Mutual empathy may represent the foundation of intimacy itself. The individuals' willingness coupled with mutual goodwill and empathy creates an atmosphere of safety and trust in which anything becomes possible. Correspondingly, the ability to trust at such deep levels is possible because of the mutual desire for deep understanding of one another.

## Cross-References

- ▶ Components of Optimal Sexual Experiences
- ▶ Lessons about Optimal Sexual Experiences from Remarkable Lovers
- ▶ Personal Contributions to Optimal Sexual Experiences

## References


Broder, M., & Goldman, A. (2004). *Secrets of sexual ecstasy*. New York: Penguin.

Castleman, M. (2004). *Great sex*. New York: Rodale.

Kleinplatz, P. J. (2006). Learning from extraordinary lovers: Lessons from the edge. *Journal of Homosexuality, 50*(3/4), 325–348.

Kleinplatz, P. J., & Ménard, A. D. (2007). Building blocks towards optimal sexuality: Constructing a conceptual model. *The Family Journal: Counseling and Therapy for Couples and Families, 15*(1), 72–78.

Kleinplatz, P. J., Ménard, A. D., Paradis, N., Campbell, M., Dalgliesh, T., Segovia, A., et al. (2009). From closet to reality: Optimal sexuality among the elderly. *The Irish Psychiatrist, 10*(1), 15–18.

Kleinplatz, P. J., Ménard, A. D., Paquet, M.-P., Paradis, N., Campbell, M., Zuccarini, D., et al. (2009). The components of optimal sexuality: A portrait of "great sex". *The Canadian Journal of Human Sexuality, 18*(1–2), 1–13.

Metz, M. E., & McCarthy, B. W. (2007). The "Good-Enough Sex" model for couple sexual satisfaction. *Sexual and Relationship Therapy, 22*, 351–362.

Metz, M. E., & McCarthy, B. W. (2012). The Good Enough Sex (GES) model: Perspective and clinical application. In P. J. Kleinplatz (Ed.), *New directions in sex therapy: Innovations and alternatives* (2nd ed., pp. 213–229). New York: Routledge.

Schnarch, D. (1997). *Passionate marriage: Love, sex, and intimacy in emotionally committed relationships*. New York: Holt.

Shaw, J. (2012). Approaching sexual potential in relationship: A reward of age and maturity. In P. J. Kleinplatz (Ed.), *New directions in sex therapy: Innovations and alternatives* (pp. 185–209). New York: Routledge.

Zilbergeld, B. (2004). *Better than ever: Love and sex at midlife*. Norwalk, CT: Crown House.


# Relational Experiences

- ▶ Relational Goods

# Relational Goods

Stefano Bartolini
Department of Economics, University of Siena, Siena, Italy

## Synonyms

Capital, social; Connections, social; Interpersonal relationships; Relational experiences; Social interactions; Social networks

## Definition

The concept of relational goods refers to one's quantity and quality of relational experience with other people. In particular relational goods concern those interpersonal relationships where the non-instrumental motivation is prevailing, i.e., those which are mostly intrinsically motivated. According to Deci (1971), "one is said to be intrinsically motivated to perform an activity when one receives no apparent reward except the activity itself." By contrast, extrinsic motives stem from outside an activity performed. Examples of relational goods that are often brought concern social and affective relationships as love, family relationships, friendship, neighborhood relations, associational activities, and many kinds of social events. However, in statistical and econometric work, in many cases it is not easy to distinguish whether a relationship is intrinsically or extrinsically motivated. For instance, friendship or a neighborhood relationship can be motivated by the pleasure of the relationship itself or by the obtainment of some reward (for instance, material rewards, exchange of favors). In this latter case the relationship is not a relational good.

## Description

From the point of view of economic theory, relational goods are a kind of local public goods which are simultaneously consumed while produced (Gui & Sugden, 2005; Uhlaner, 1989). Relational goods are local public goods since non-excludability (nobody can be excluded from access to a good) and non-rivalry (one's increase in consumption of a good does not diminish others' consumption of the same good) are limited to those having access to the relational good. Moreover, "they are simultaneously consumed and produced since participating to them is both an act of production (my presence increases the value of the good) and consumption (I enjoy it while being present)" (Becchetti, Pelloni, & Rossetti, 2008).

Relational goods are a component of the broader concept of social capital. Indeed the latter refers to the relationships between people and institutions, to that among institutions and to that among people. It is only this latter component of social capital – the relationships among people – which is considered within the concept of relational goods.

Research, particularly the economic one, have focused on the role of relational goods in determining the QOL. More precisely, studies concentrated on the linkage between relational good and well-being. Indeed, econometric work on surveys administered on nations' representative samples showed a robust correlation between several measures of relational goods and subjective well-being (see the pioneering studies by Helliwell (2001, 2006) and Helliwell and Putnam (2004)), although they do not use the term relational goods (see also Bruni and Stanca (2008), Becchetti et al. (2008), Powdthavee (2008)). Moreover, changes over time in relational good proved to be a very good predictor of the trends of subjective well-being in the long and the medium run (Bartolini, Bilancini, Pugno, forthcoming, Bartolini & Sarracino, 2011). Finally, Becchetti, Giachin Ricca, and Pelloni (2009) provided a causal analysis showing that social capital has a strong impact on SWB.

These economic papers echo a large psychological literature emphasizing the strong association between physical and mental well-being and the quality of relational experience. For instance, according to Carr (2004), "people with large social support networks and stronger social bonds with members of their

networks have better physical and mental health, fewer illnesses and less depression, recover more rapidly from physical illness and psychological problems, and have a lower risk of death."

The decline in relational goods has been proposed as one of the explanations of the Easterlin paradox, i.e., the evidence that people do not become happier when a country's income grows (Easterlin, 1974, 1995). Indeed, Bartolini et al. (forthcoming) have shown that the decline in relational goods is the main predictor of the decline in the average happiness of Americans over the last 30 years.

## Cross-References

► Subjective Well-Being

## References

Bartolini, S., & Sarracino, F. (2011). Happy for How Long? Social Capital, GDP and Well-Being over Time, Quaderni del Dipartimento di Economia Politica, Università degli Studi di Siena, N. 621.

Bartolini, S., Bilancini, E., & Pugno, M. (2013). Did the decline in social connections depress Americans' happiness? *Social Indicators Research*, 110, (3) pp. 1033-1059.

Becchetti, L., Giachin Ricca, E., & Pelloni, A. (2009). The 60es turnaround as a test on the causal relationship between sociability and happiness. Econometica. Working papers wp07.

Becchetti, L., Pelloni, A., & Rossetti, F. (2008). Relational goods, sociability and happiness. *Kyklos, 61*(3), 343–363.

Bruni, L., & Stanca, L. (2008). Watching alone. Happiness, relational goods and television. *Journal of Economic Behavior and Organization, 65*(3–4), 506–528.

Carr, A. (2004). *Positive psychology*. London: Brunner-Routledge.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*, 105–115.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David & W. R. Melvin (Eds.), *Nations and households in economic growth* (pp. 98–125). Palo Alto, CA: Stanford University Press.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization, 27*, 35–47.

Gui, B., & Sugden, R. (2005). *Economics and social interaction*. Cambridge: Cambridge University Press.

Helliwell, J. F. (2001). Social capital, the economy and wellbeing. In *The review of economic performance: The longest decade: Canada in the 1990s*. Ottawa, ON: Centre for the Study of Living Standards.

Helliwell, J. F. (2006, March). Well-being, social capital and public policy: hat's new? *The Economic Journal, 116*, C34–C45.

Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. *Philosophical Transactions, 359*(1449), 1435–1446.

Powdthavee, N. (2008). Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships. *Journal of Socio-Economics, 37*(4), 1459–1480.

Uhlaner, C. J. (1989). Relational goods and participation: Incorporating sociability into a theory of rational action. *Public Choice, 62*(3), 253–285.

## Relational Self

► Independent/Interdependent Self

## Relational Sex Motives

► Relationship Contingency and Sexual Satisfaction

## Relations with God

► Spiritual Struggles

## Relationship Conflict and Health

► Marital Conflict and Health

# Relationship Contingency and Sexual Satisfaction

Corinne A. Moss-Racusin
Yale University, Skidmore College,
Saratoga Springs, NY, USA

## Synonyms

Appetitive intimacy sex motives; Approval sex motives; Aversely driven approval sex motives; Contingencies of self-worth and sexual satisfaction; Deriving self-worth from romantic relationships; Intimacy sex motives; Preference, sexual; Relational sex motives

## Definition

Relationship contingency (RC) refers to deriving self-worth from dating relationships and has implications for women's sexual motivation and sexual satisfaction within romantic relationships. Because RC enhances goals to sustain relationships to maintain positive self-worth, RC women tend to engage in sexual behavior to maintain and enhance their relationships (i.e., they demonstrate *relational sex motives*). However, different types of relational sex motives predict different sexual outcomes for women. Whereas having sex to improve intimacy with one's partner (i.e., *intimacy motives*) is associated with greater sexual autonomy and sexual satisfaction, having sex to gain approval from one's partner (i.e., *approval motives*) is associated with reduced sexual autonomy and satisfaction.

## Description

When people derive their self-esteem from specific domains (i.e., their self-worth is based on excelling in particular areas of life), they may be said to be contingent upon those domains. Contingencies of self-worth theory (Crocker & Knight, 2005) describe the complicated relationship between people's self-esteem and the specific domains from which they derive it. Of importance, domain contingencies often increase people's effort and motivation to sustain success in these particular domains, in order to maintain self-esteem and feelings of positive self-worth. For example, people who are contingent upon academic achievement tend to spend extra time studying, whereas those who are contingent upon their physical appearance may spend extra time engaged in exercise or tending to their appearance (Crocker & Luhtanen, 2003). In this way, contingencies of self-worth (CSWs) can have positive outcomes, in that they enhance goals and motivations within contingent domains. However, CSWs have also been shown to have negative consequences, in that they can undermine self-confidence, interpersonal relationships, and learning, due to the fact that failures in contingent domains have particularly high consequences for well-being (Crocker & Knight, 2005). Research by Sanchez, Moss-Racusin, Phelan, and Crocker (2011) on the role of CSWs in determining motivation in sexual relationships has revealed that being relationship contingent affects when and why people engage in sexual behavior with their partner.

Romantic relationships are one example of a domain from which people can base their self-esteem. When considering possible consequences, being relationship contingent (relative to academic or even appearance contingent) may have particularly mixed outcomes. Because their self-worth is tied to their interpersonal connection to their primary romantic relationship partner, individuals who are relationship contingent (RC) experience greater emotional highs and lows as a result of their romantic relationships. That is, RC individuals exhibit greater jumps in self-esteem after positive romantic relationship events (such as an engagement or marriage), as well as more severe drops in self-esteem and heightened distress following negative relationship events (such as a large fight or divorce) relative to people who are not RC (Knee, Canavello, & Blush, 2008). Similarly, high-RC people are particularly attentive to their partners needs (Park, Sanchez, & Brynildsen, 2011) and are more committed to their relationships

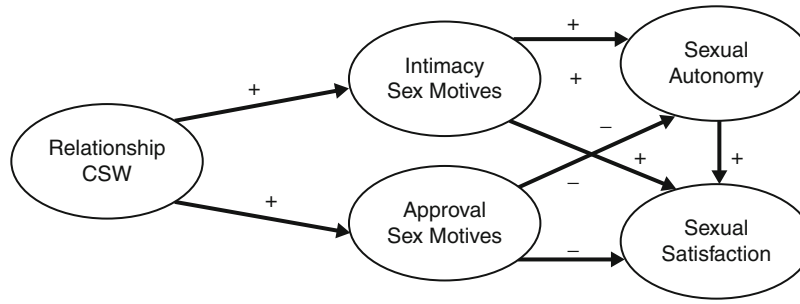
(Knee et al., 2008) than low-RC individuals. However, high RC is also associated with the obsessive pursuit of former partners after a relationship has ended (Park et al., 2011), and there is also evidence that high levels of relationship commitment do not necessarily translate into more satisfying relationships (Knee et al., 2008). Indeed, high-RC individuals' fear of being without a romantic relationship appears to motivate them to remain in unhealthy relationships for excessive periods of time (Sanchez, Good, Kwang, & Saltzman, 2008). Thus, evidence suggests that being RC has both positive and negative outcomes for romantic relationships.

Turning to more specific outcomes, it is critical to consider the impact of relationship contingencies on sexual behavior and satisfaction rather than just broader relationship effects. Because being RC makes people more likely to engage in behaviors aimed at preserving their relationship, it is also likely to have implications for sexual behavior. That is, because experiencing sexual satisfaction is associated with relationship satisfaction (Christopher & Sprecher, 2000), people who are RC may be more motivated to have sex as a means of preserving their romantic relationships, rather than simply experiencing sexual pleasure. Theories of sexual motivation refer to these types of goals as *relational sex motives*, that is, motivations to engage in sexual behavior with the broader aim of creating positive outcomes in the relationship.

However, within the larger framework of relationship sex motives, there are two distinct types of motives likely to stem from RC, with drastically different implications for sexual satisfaction (e.g., Impett, Gable, & Peplau, 2005). Though both can potentially address the goal of having sex to strengthen the relationship, the differing motivational underpinnings result in either positive or negative outcomes in terms of pleasure and sexual functioning. More specifically, *approval sex motives* involve engaging in sexual behavior in order to avoid negative outcomes such as rejection and disapproval from a romantic relationship partner. As such, approval sex motives are focused on the avoidance of negative relationship outcomes and reflect aversely motivated behaviors. In contrast, *intimacy sex motives* involve having sex to enhance intimacy and closeness in relationships, are focused on creating positive relationship outcomes, and thus reflect appetitive behaviors.

In sum, the outcomes of relational sex motives differ as a result of whether the behavior is the result of the desire to please a partner (aversely driven approval sex motives) or the desire to create and foster closeness (appetitive intimacy sex motives). Research on gender role conformity has uncovered a similar pattern, whereby personally investing in societal gender ideals was positively related to self-esteem when gender-conforming behaviors were appetitive and freely chosen (i.e., autonomous) but negative when they were engaged in aversely as a result of pressure from others or the situation (i.e., pressured; Good & Sanchez, 2010). In this case, engaging in sex primarily to gain the approval of a partner dampens the expression of one's own sexual desires and preferences. In turn, approval sex motives decrease feelings of sexual autonomy (or independence) and increase unwanted sexual behaviors, resulting in depressed psychological well-being and relationship satisfaction over time (Impett et al., 2005).

Research has indicated that this negative chain of events may be particularly true for women, who are more likely than men to participate in the type of unwanted passive sexual behavior which undermines sexual functioning and satisfaction (Sanchez, Kiefer, & Ybarra, 2006). Conversely, women who experience intimacy sex motives and engage in sexual behavior in order to enhance closeness with their partners likely experience more sexual autonomy, to the benefit of their sexual satisfaction. Simply put, women who are motivated to engage in sexual behavior in order to get closer to their partners are more likely to communicate their sexual desires and preferences and, as a result, are more likely to have them met. Moreover, women's sexual submissiveness is negatively associated with both their own and their male partner's sexual satisfaction when it is inconsistent with their own sexual preferences (i.e., when their sexual autonomy is low; Sanchez, Phelan, Moss-Racusin, & Good, 2012). For these reasons, women are




**Relationship Contingency and Sexual Satisfaction,**
**Fig. 1** Sanchez et al. (2011) model of relationship contingency, sexual motivations, and sexual satisfaction

a particularly important group to examine when it comes to understanding the impact of RC on sexual satisfaction.

Sanchez et al. (2011) were the first to demonstrate these complex relationships between relationship women's contingencies of self-worth (i.e., being relationship contingent), relational sex motives (approval vs. intimacy), and the outcomes of sexual autonomy and sexual satisfaction in a single model (see Fig. 1). Specific predictions derived from their model were as follows:

Hypothesis 1: Women's relationship contingency predicts greater relational sexual motivation (e.g., higher intimacy sex motives and approval sex motives).

Hypothesis 2: Relational sexual motivations predict sexual autonomy and satisfaction, such that intimacy sex motives predict greater sexual autonomy and satisfaction, while approval sex motives predict lower sexual autonomy and satisfaction.

Hypothesis 3: Greater sexual autonomy in turn predicts greater sexual satisfaction.

Exploratory analyses included in the model: Lastly, exploratory analyses were conducted to examine whether relationship length and approval contingency affect whether relationship contingency is associated with intimacy sex motives or approval sex motives.

Results of structural equation modeling analyses supported each of the hypotheses. More specifically, RC was related to heightened relational sex motives via two distinct paths. On the one hand, RC predicted increased intimacy sex motives, which were related to enhanced sexual autonomy and satisfaction. On the other hand, RC was also linked to higher levels of approval sex motives, which reduced sexual autonomy and satisfaction. Of importance, the analyses indicated that the gender of women's partners did not impact results. That is, the model provided an equally good fit for women in same-sex and opposite sex relationships, suggesting that the processes operate similarly for women who self-identify as straight and as lesbian or bisexual. However, it is unknown whether the processes would operate similarly for men. Future research should seek to generalize the model by testing male participants.

In terms of the exploratory analyses, there was preliminary evidence that relationship length and people's general predisposition to seek approval from others (i.e., approval contingency) affect whether RC predicts intimacy sex motives or approval sex motives. That is, when people are generally less focused on seeking approval in all of their relationships (i.e., when they are lower on approval contingency), RC predicts greater intimacy sex motives. Additionally, when people are in the beginning stages of their romantic relationships, RC predicts heightened approval sex motives. These exploratory results help shed light on the factors that determine whether relational sex motives result in intimacy or approval sex motives. However, it is clear that more work should focus on further illuminating these pathways, particularly because they lead to such divergent outcomes for sexual autonomy and satisfaction.

This model is critical in that it directly examines the association between relationship contingencies of self-worth and sexual satisfaction. Moreover, it identifies two unique processes by

which RC can result in either positive or negative outcomes for sexual autonomy and satisfaction. When RC is associated with appetitive intimacy sex motives focused on enhancing and sustaining closeness with a romantic partner, women experience greater levels of sexual autonomy, to the benefit of sexual satisfaction. In contrast, when RC leads to adversely driven approval sex motives aimed at avoiding negative relationship outcomes, women's sexual autonomy and ultimately their levels of sexual satisfaction suffer as a result. Thus, the model tested by Sanchez et al. (2011) illuminates the complex and bivalenced relationship between women's relationship contingencies of self-worth and their sexual satisfaction.

## References


Christopher, F. S., & Sprecher, S. (2000). Sexuality in marriage, dating and other relationships: A decade review. *Journal of Marriage and the Family, 62*, 999–1017.

Crocker, J., & Knight, K. M. (2005). Contingencies of self-worth. *Current Directions in Psychological Science, 14*, 200–203.

Crocker, J., & Luhtanen, R. K. (2003). Level of self-esteem and contingencies of self-worth: Unique effects on academic, social, and financial problems in college freshmen. *Personality and Social Psychology Bulletin, 29*, 701–712.

Good, J. J., & Sanchez, D. T. (2010). Doing gender for different reasons: Why gender norm conformity positively and negatively predicts self-esteem. *Psychology of Women Quarterly, 34*, 203–214.

Impett, E. A., Gable, S. L., & Peplau, L. A. (2005). Giving up and giving in: The costs and benefits of daily sacrifice in intimate relationships. *Journal of Personality and Social Psychology, 89*, 327–344.

Knee, C. R., Canavello, A., & Bush, A. L. (2008). Relationship contingent self-esteem and the ups and downs of romantic relationships. *Journal of Personality and Social Psychology, 95*, 608–627.

Park, L. E., Sanchez, D. T., & Brynildsen, K. (2011). Maladaptive responses to relationship dissolution: The role of relationship contingent self-worth. *Journal of Applied Social Psychology, 7*, 1749–1773.

Sanchez, D. T., Good, J., Kwang, T., & Saltzman, E. (2008). When finding a mate becomes urgent: Why relationship contingency predicts men's and women's body shame. *Social Psychology, 39*, 90–102.

Sanchez, D. T., Kiefer, A., & Ybarra, O. (2006). Sexual submissiveness in women: Costs for autonomy. *Personality and Social Psychology Bulletin, 32*, 512–524.

Sanchez, D. T., Moss-Racusin, C. A., Phelan, J. E., & Crocker, J. (2011). Relationship contingency and sexual motivation in women: Implications for sexual satisfaction. *Archives of Sexual Behavior, 40*, 99–110.

Sanchez, D. T., Phelan, J. E., Moss-Racusin, C. A., & Good, J. J. (2012). A gender role motivation model of women's sexually submissive behavior and satisfaction in couples. *Personality and Social Psychology Bulletin, 38*, 528–539.


## Relationship Interdependence and Health

▶ Marital Conflict and Health

## Relationship Quality

▶ Marital Well-being Measures

## Relationship Quality and Commitment in Norway and Sweden

Kenneth Aarskaug Wiik
Research Department, Statistics Norway, Oslo, Norway

### Definition

Differences in relationship quality by union type have been the focus of numerous, mostly US, studies, of which the majority have concluded that married individuals are more satisfied with their relationship than cohabitors (e.g., Brown, 2003, 2004; Brown & Booth, 1996; Stanley, Whitton, & Markman, 2004). Comparing those who cohabited with their spouse prior to marriage with those who married directly without cohabiting first and those who were currently cohabiting, Nock (1995) found that cohabitors were less happy with their relationships than both groups of married individuals.

Interpersonal commitment, referring to the level of commitment to a specific partner or union, on the other hand, has received little attention in the study of cohabitation patterns, despite

the fact that it seems to be strongly related to relationship quality, which in turn is associated with various relationship outcomes, such as dissolution risks and interpartner ▶ violence (Amato, 2007; Stanley et al., 2004). Defining commitment as the perceived costs of exiting a union, Nock (1995) reported lower levels of commitment among cohabiting relative to married individuals in the USA. Further, comparing currently married individuals with cohabitors without definite marriage plans, Stanley et al. (2004) found that the first group was significantly more dedicated to their relationships (i.e., a desire to prioritize the relationship), even after controlling for satisfaction levels.

As cohabitation serves as a prelude to marriage for many cohabiting couples, some studies take into account cohabitors' marriage plans. For instance, Brown and Booth (1996) found that the relationships of cohabitors with marriage plans were not qualitatively different from those of marrieds. Therefore, analyses showing poorer relationship quality among cohabitors could reflect the lack of a control for marriage plans, which, in turn, is related to commitment. Similarly, there seems to be no or only small differences in relationship quality between engaged cohabitors and those who are already married (Kline et al., 2004; Rhoades, Stanley, & Markman, 2009).

Studies of union commitment or satisfaction seem to be rare in the Scandinavian countries. One recent exception is Hansen, Moum, and Shapiro (2007), who analyzed relational well-being in Norway. They found that cohabitors reported lower levels of relationship happiness compared with married individuals. Wiik, Bernhardt, and Noack (2009) investigated the relation between cohabitation and commitment and union quality in Norway and Sweden, two countries where cohabitation is more widespread than in most other countries and where the living arrangement is socially acceptable and equal to marriage in terms of public policy. They used data from two comparable surveys on married and cohabiting men and women aged 25–35 conducted in 2003 ($N = 2{,}923$). In this study, three dimensions of commitment and relationship quality were considered: relationship seriousness, satisfaction with the relationship, and breakup considerations. They also took into account whether the cohabitors had concrete marriage intentions or not.

## Description

Wiik, Bernhardt, and Noack (2009) measured relationship seriousness by asking respondents to rate the seriousness of their present partnership. For purposes of comparison across the two countries answers were dichotomized according to whether respondents viewed their unions as more (1) or less (0) serious. *Relationship satisfaction* was tapped by asking respondents how satisfied they were with their unions. This variable was regrouped as a dummy indicating whether respondents were very (1) or moderately to less satisfied (0) with their partnership. They chose this method because the number of respondents rating their relationship as not satisfying was low in both surveys (only about 6 % had a value 3 or lower). The last dependent variable was made by utilizing a question asking respondents whether they or their current partner had *considered breaking up the union* in the year preceding the survey ($1 = yes$, $0 = no$).

Table 1 shows that for each of the three outcomes, there were no major differences in their distributions between married respondents and cohabitors with marriage plans (planned to marry their current partner within two years), although a significantly higher share of the latter group was satisfied with their union. Cohabitors without marriage plans, on the other hand, were less serious and satisfied and more often had thought of ending the relationship than marrieds.

Results from three separate logistic regression models of the likelihood of reporting a serious relationship, being satisfied with one's relationship, and having considered ending the current union are presented in Table 2. These results confirm that married respondents scored higher than cohabitors without marriage intentions on relationship commitment and satisfaction, but did not score significantly higher than cohabitors with marriage plans. This holds true even when

controlling for age, age homogamy, gender, union duration, presence of common children or stepchildren, birth plans, respondents and their partner's annual income, home ownership, educational level, religiosity, and country (Wiik, Bernhardt, & Noack, 2009).

More precisely, cohabiting respondents without marriage intentions were significantly less likely to be serious about their current relationship than married respondents. Comparing the odds ratio estimates for the two groups of respondents, cohabitors without marriage plans were only 36 % as likely to be serious about their current union compared with marrieds, net of the other variables included. Table 2 further indicates that cohabitors without marriage intensions were significantly less satisfied with their present union than was the case for those married. Net of the other variables included, this reduction in the likelihood of being most satisfied corresponded to 60 % for cohabitors without plans to marry, relative to married individuals.

To be sure, research has found that higher levels of commitment might explain why married individuals are more satisfied with their unions than cohabitors (e.g., Nock, 1995). Conversely, more satisfaction among those married could modify some of the differences between the union forms in commitment (Stanley et al., 2004). To test whether the differences between the union types in satisfaction were mediated by commitment and vice versa, union seriousness and satisfaction were included as covariates in two supplementary models of satisfaction and seriousness, respectively. Although displaying significantly positive effects on the two outcomes, the negative relation between cohabiting with no plans to marry and satisfaction and seriousness remained stable in these additional models (Wiik, Bernhardt, & Noack, 2009).

From the last model presented in Table 2 it is evident that cohabitors, their partners, or both had an elevated risk of considering splitting up than was the case for those married. This increase in breakup plans was, however, only statistically significant for cohabitors without marriage plans.

**Relationship Quality and Commitment in Norway and Sweden, Table 1** Descriptive statistics by union type, percent ($N$ = 2,923) (Copyright National Council on Family Relations, 2009)

| | Married (%) | Cohabiting, intend to marry this partner (%) | Cohabiting, no intent to marry this partner (%) |
|---|---|---|---|
| Dependent variables | | | |
| Serious[a] | 91.0 | 91.2 | 77.3* |
| Satisfied[a] | 69.7 | 78.6* | 58.6* |
| Breakup plans[a] | 12.1 | 13.9 | 28.4* |
| *n* | 1,326 | 310 | 1,287 |

Source: Wiik, Bernhardt, and Noack (2009)
[a]0 = moderately to least serious, satisfied, or no breakup plans; 1 = most serious, satisfied, or have breakup plans
*Difference between cohabitors and married respondents is statistically significant at $p < .05$

**Relationship Quality and Commitment in Norway and Sweden, Table 2** Results from three logistic regression models of (a) relationship seriousness ($N$ = 2,876), (b) satisfaction ($N$ = 2,903), and (c) breakup plans ($N$ = 2,900) (Copyright National Council on Family Relations, 2009)

| | Seriousness | | Satisfaction | | Breakup plans | |
|---|---|---|---|---|---|---|
| | B (SE B) | OR | B (SE B) | OR | B (SE B) | OR |
| Married (=reference) | | | | | | |
| Cohabiting, plan to marry | −0.02 (0.24) | 0.98 | 0.03 (0.16) | 1.03 | 0.21 (0.20) | 1.23 |
| Cohabiting, no marriage plans | −1.03*** (0.14) | 0.36 | −0.92*** (0.10) | 0.40 | 1.07*** (0.12) | 2.91 |
| % | 84.7 | | 65.8 | | 19.5 | |

Source: Wiik, Bernhardt, and Noack (2009)
Note: 0 = moderately to least serious, satisfied, or no breakup plans; 1 = most serious, satisfied, or have breakup plans. *OR* odds ratio
Models controlled for age, age homogamy, gender, union duration, presence of common children or stepchildren, birth plans, respondents and their partner's annual income, home ownership, educational level, religiosity, and country
***$p < .001$

Controlling for relevant characteristics of the respondents themselves and their partners, this relative rise in the risk of breakup plans was nearly three times as high for cohabitors without intentions to marry. As lower levels of relationship satisfaction and seriousness could account for the elevated risk of considering a breakup among cohabitors without marriage plans, seriousness and satisfaction were included in an additional regression model of breakup plans. The positive relation between cohabiting without intent to marry and having breakup plans remained stable in this alternative model.

Possible interactions between union type and gender, the number of common children, union duration, and country were also tested. First, the product term representing the interaction between gender and cohabitation with intention to marry was positive and statistically significant ($p < .05$). Female cohabitors intending to marry their partners were, in other words, more serious than their male counterparts. Also, there was a significant positive interaction between type of union and the number of common children on relationship seriousness, implying that both types of cohabitors with one common child were more serious than childless cohabitors. Further, the negative main effect of cohabitation relative to marriage on relationship seriousness decreased with the duration of the union. Last, cohabitors with marriage plans having two or more children were less satisfied than the childless cohabitors (Wiik, Bernhardt, & Noack, 2009).

## Discussion

Even in Scandinavia, a region that has been at the vanguard of changes in family structures, relationship commitment and quality vary with type of union. Results from the study by Wiik, Bernhardt, and Noack (2009) confirmed that married respondents were significantly more committed and reported higher levels of relationship quality than cohabitors without marriage plans on all three dimensions considered (relationship satisfaction, commitment, and breakup plans). There were, however, considerable differences between cohabitors with and without marriage intentions. The former group was far more serious, satisfied, and less often had dissolution plans than the latter. These results held true controlling for a range of possible confounders.

The contention that cohabitors are a heterogeneous group was further supported by the fact that the association between union type and relationship assessments varied with other variables. First, female cohabitors with marriage intentions were more serious about the union than their male counterparts, and the difference between cohabitors and those married in relationship seriousness decreased with duration. Next, cohabitors with one common child were more serious about their current relationship than the childless, whereas marriage-prone cohabitors having more than one child were less satisfied than those without children.

The higher degree of commitment and satisfaction among married individuals can be a consequence of the marriage itself and the norms and values associated with the institution of marriage. The fact that cohabitors in the two countries continue to get married suggests that, for some reason or another, marriage is still the preferred form of living together, even in societies where cohabitation is a completely accepted social and civil status even for having and rearing children. Although marriage-like cohabitations (i.e., those who have common children or have lived together for some years) are essentially equal to married couples in terms of public policy, the union forms continue to be diverse in the area of private law. It is also worth noticing the differences between entering cohabitation and marriage. Not only the wedding ceremony itself but several rituals and practices remain reserved for marriage.

On the other hand, comparing cohabitors and those married, we will inevitably be confronted with the issue of selection, i.e., whether married couples at the start were more committed and satisfied. Considering that cohabitors with marriage intentions reported higher levels of relationship commitment and satisfaction than other cohabitors, it seems plausible that there is at least a certain selection effect and that the degree of satisfaction and commitment are important factors that triggers the transformation of cohabiting relationships into marriages.

## References


Amato, P. R. (2007). Transformative processes in marriage: Some thoughts from a sociologist. *Journal of Marriage and Family, 69*, 305–309.

Brown, S. L. (2003). Relationship quality dynamics of cohabiting unions. *Journal of Family Issues, 24*, 583–601.

Brown, S. L. (2004). Moving from cohabitation to marriage: Effects on relationship quality. *Socials Science Research, 33*, 1–19.

Brown, S. L., & Booth, A. (1996). Cohabitation versus marriage: A comparison of relationship quality. *Journal of Marriage and the Family, 58*, 668–678.

Hansen, T., Moum, T., & Shapiro, A. (2007). Relational and individual well-being among cohabitors and married individuals in midlife: Recent trends from Norway. *Journal of Family Issues, 28*, 910–933.

Kline, G. H., Stanley, S. M., Markman, H. J., Olmos-Gallo, P. A., Peters, M. S., Whitton, S. W., et al. (2004). Timing is everything: Pre-engagement cohabitation and the increased risk for poor marital outcomes. *Journal of Family Psychology, 18*, 311–318.

Nock, S. L. (1995). A comparison of marriages and cohabiting relationships. *Journal of Family Issues, 16*, 53–76.

Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2009). The pre-engagement cohabitation effect: A replication and extension of previous findings. *Journal of Family Psychology, 23*, 107–111.

Stanley, S. M., Whitton, S. W., & Markman, H. J. (2004). Maybe I do. Interpersonal commitment and premarital or nonmarital cohabitation. *Journal of Family Issues, 25*, 496–519.

Wiik, K. A., Bernhardt, E., & Noack, T. (2009). A Study of commitment and relationship quality in Sweden and Norway. *Journal of Marriage and Family, 71*, 465–477.


# Relationship Satisfaction


Renske Keizer
Sociology, Faculty of Social Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands


## Synonyms

Dyadic satisfaction; Marital satisfaction; Satisfaction with living partner; Spousal satisfaction; Terminology concerning relationship satisfaction

Many terms have been used in the literature to refer to relationship satisfaction, i.e., relationship quality, relationship adjustment, spousal satisfaction, and relationship happiness. Many of these terms have been used interchangeably, and there is still little agreement in the literature on the use of a single term (see Kluwer, 2010). Furthermore, although the majority of studies that focus on marital satisfaction do indeed only incorporate couples that are married, many studies that use the term "marriage" and "marital" do not mean to ignore or exclude unmarried couples.

## Definition

Relationship satisfaction is the subjective evaluation of one's relationship. Relationship satisfaction is not a property of a relationship; it is a subjective experience and opinion. As such, members of the same couple may differ in how satisfied they are with their relationship.

## Description

### Relevance of the Topic

Relationship satisfaction has been the subject of research for many decades.

According to Bradbury, Fincham, and Beach (2000) who conducted one of the most recent reviews on relationship satisfaction, the rationale for studying relationship satisfaction "stems from its centrality in individual and family well-being, from the benefits that accrue to society when strong relationships are formed and maintained, and from the need to develop empirically defensible interventions for couples that prevent or alleviate relationship distress and breakups and divorce" (p. 964).

### Focus of the Current Entry

Most scholars have investigated relationship satisfaction in one specific relationship, namely, marriage: "marital satisfaction." In the current entry, to avoid confusion, I will use the term relationship satisfaction to refer to satisfaction

within couples, irrespective of whether these couples are living apart together, living together unmarried, or married.

The research described below concerns relationship satisfaction within heterosexual couples. For detailed information on relationship satisfaction within lesbian, gay, bisexual, and transgender relationships, see the articles of Biblarz and Savci (2010), Kurdek (2003, 2004), and Patterson (2000).

On average, individuals report to be very satisfied with their relationship. This finding is of course related to attrition: individuals who were very unsatisfied with their relationship are more likely to have already broken up with their partner.

## Differences in Relationship Satisfaction

### Differences by Relationship Status

Several studies have revealed that relationship satisfaction varies by relationship status. In general, studies show that unmarried cohabitors and repartnered couples in general are less satisfied with their relationship compared with those married. Furthermore, the literature showed that transitions into marriage are shown to benefit relationship satisfaction and that couples with more positive relationship assessments were more likely to enter marriage.

**Differences Across Cultures** Evolutionary research and theory have emphasized that there are ubiquitous partnership arrangements in relationships. However, scholars have also showed that interpersonal relationships are heavily guided by norms, customs, and expectations that are derived from culture. In particular, satisfaction with one's spouse may largely depend on the degree to which a relationship fulfills culturally determined expectations and obligations (Lucas et al., 2008). For example, playful teasing within the context of romantic relationships may be viewed more positively by Eastern than Western cultures (Campos, Keltner, Beck, Gonzaga, & John, 2007). In general, studies have shown that average relationship satisfaction is higher in Western than in non-Western countries.

## Determinants of Relationship Satisfaction

According to an influential review of more than 100 longitudinal studies of marriage (Karney & Bradbury, 1995), the following determinants are involved:

1. Spouses' background and characteristics (e.g., personality, parental divorce)
2. Life stressors and transitions while in relationship (e.g., stress related to work and health, the transition to parenthood)
3. Interactional/couple processes (e.g., emotions and communication skills that spouses display while adapting to each other and the stress they confront)

### Spouses' Background and Characteristics (e.g., Personality, Parental Divorce)

Among these factors, personality, family background, religion, ethnicity, personal history, early experiences, and personal preferences are most often studied. I will elaborate on the first three factors here.

Several studies reveal that broad personality constructs are associated with relationship satisfaction. People with low levels of neuroticism, high levels of agreeableness, and high levels of extraversion are more likely to report high levels of relationship satisfaction (e.g., Kurdek, 1997).

More and more scholars also come to recognize the importance of intergenerational associations when it comes to relationship satisfaction. Scholars have shown that variation in relationship outcomes can be linked to relationship functioning of the spouses' parents. For example, parents' reports of marital quality are related with their offspring's marital quality assessed more than 10 years later. These findings lend support to the widely held view that the family of origin is a primary setting in which children and adolescents learn (mal) adaptive interpersonal repertoires (Bradbury & Karney, 2004).

Religious involvement appears to have a positive impact on relationship satisfaction, mainly on marital satisfaction, and this effect appears to be particularly salient among lower income couples.

### Life Stressors and Transitions While in Relationship (e.g., Stress Related to Work and Health, the Transition to Parenthood)

Among these factors, the transition to parenthood and paid work and household labor are most often studied.

A very large part of the literature on relationship satisfaction scrutinizes how children and entering parenthood affects relationship satisfaction. The transition to parenthood literature has consistently documented a decline in relationship satisfaction for parents, with this decline being greater for women than men, especially when the children are young (see for reviews of Kluwer, 2010 and Twenge, Campbell, & Foster, 2003). Researchers have examined how couples negotiate the transition to parenthood. These studies have shown that there is large variability in how couples' relationship satisfaction is affected by the arrival of the firstborn. Most scholars also draw on the three domains of influence mentioned above: spouses' background and characteristics, life stressors and transitions while in relationship, and interactional/couple processes. Kluwer (2010) shows in her review that marital conflict increases across the transition to parenthood and that marital interactions become increasingly negative and decreasingly positive. Furthermore, her work reveals that adaptive processes such as relationship maintenance, emotional responsiveness, and spousal support decrease across the transition to parenthood. Regarding spouses' background and characteristics, life stressors, and life transitions, Kluwer's review shows that depression, the birth of a daughter, insecure attachment, and incongruence between work and division of labor status versus personal preferences render couples more vulnerable to negative changes in relationship satisfaction across the transition to parenthood.

Next to the transition to parenthood, economic, household, and work-related stressors comprise the largest body of research on the impact of life stressors on relationship satisfaction (i.e., job quality, work stress, work overload, night shifts, division of household labor). Some studies showed that a large number of work hours and work stress are associated with little marital interaction and low marital satisfaction. Other studies showed that working nonstandard work schedules negatively affect relationship satisfaction and increase the risk of divorce. Regarding household labor, few studies found direct effect on relationship satisfaction. Most studies found that household labor affects relationship satisfaction via the perceived fairness of, and the satisfaction with, the division of household labor.

Several studies have examined the impact of other life stressors on relationship satisfaction, such as traumatic events, child illness or child death, and illnesses of one's partner. Many of these studies document not only the diverse ways that couples adapt to these extreme difficulties but also the remarkable resilience that they display when doing so (Bradbury, Fincham, & Beach, 2000).

### Interactional/Couple Processes (e.g., Emotions and Communication Skills That Spouses Display While Adapting to Each Other and the Stress They Confront)

Scholars that have conducted behavioral analyses of interactions between spouses identify specific affective exchanges (e.g., high levels of anger and contempt, low levels of affection and support) that predict lower levels of relationship satisfaction. The three most commonly assessed interactional processes are the couple's communication skills, positive and negative emotions and social/partner support (Fincham & Beach, 2010).

## Linking Contextual Factors with Interactional Processes

Karney and Bradbury (1997) showed that personality factors (e.g., negative affectivity agreeableness) predict levels of relationship satisfaction (but not change in these levels), whereas observed behavioral variables predict the degree to which relationship satisfaction deteriorates (but not the level of relationship functioning). These findings suggest that an individual's characteristics and background may contribute to relationship satisfaction by lowering or raising the overall level of satisfaction that a given spouse reports, whereas behaviors and expressions

exchanged between partners may contribute to the degree to which satisfaction changes around that level.

## Why Does Satisfaction Decline in Relationships?

One general finding in the literature is that relationship satisfaction is not a stable attitude. Most studies conducted in Western societies reveal that relationship satisfaction decreases after the first years of being together.

Scholars have offered several explanations for these changes in relationship satisfaction. The first concerns *set-point theory*, which takes normal rebounds as a starting point, and argues that relationship satisfaction is a rather stable judgment over a long period of time. Set-point theory, although mostly applied to SWB in general, emphasizes the impact of personality on satisfaction. This theory is also known as the dynamic equilibrium theory (Haedey & Wearing, 1989). The theory is derived from Helson's adaptation-level theory (1964) and assumes that everyone has a certain baseline or equilibrium level of well-being, determined by genes and personality.

The event of falling in love and getting married raises an individual's feelings of relationship. According to set-point theory, these temporary higher levels of relationship satisfaction will return to the pre-event level of relationship satisfaction over time. In the marital satisfaction literature, this phenomenon is called "the honeymoon-is-over effect" (Kurdek, 1998) or the "typical honeymoon then years of blandness pattern" (Aron, Norman, Aron, & Lewandowski, 2002). This quick and complete return to the pre-event level is called the adaptation effect. As people begin to return to their set levels of relationship satisfaction, people report decreases in relationship satisfaction (i.e., in comparison to the peak of getting married).

Two other models have been proposed to explain how relationship satisfaction, in general, deteriorates over time (e.g., Huston, Caughlin, Houts, Smith, & George, 2001; Kurdek, 2002). In contrast to the set-point theory, which is most often used for general feelings of well-being, the *emergent distress model* mainly describes changes in relationship satisfaction and proposes that relationship distress is the result of changes in relationship functioning. Couples start their relationship with idealized notions of each other and their relationship. After a while, partners become more realistic and more aware of the limitations of their partner and relationship, leading to increased distress and decreased relationship satisfaction. In contrast to the emergent distress model, *the enduring dynamics model* poses that individuals begin their relationship while they are aware of their partner's shortcomings and that relationship distress stems from problems that existed at the beginning of the relationship, such as incompatibility in personalities.

In general, there is now consensus in the literature that relationships on average are viewed as less satisfying as time passes, although there is some debate over whether these changes are primarily linear or nonlinear in form and whether rebounds may occur. The set-point theory, the emergent distress model, and the enduring dynamics model have in common that they can be used to explain *why* relationship satisfaction *in general* declines as time passes.

However, most scholars agree that a focus on aggregated and average patterns of change in relationship satisfaction masks the fact that for some couples, relationship satisfaction may decline rapidly, while for others their relationship satisfaction does not decline much at all. This recognition led to a focus on the conditions under which individual's and couples' relationship satisfaction changes. Since that moment, several (longitudinal) studies have focused on understanding the variability in people's patterns of change in relationship satisfaction (e.g., Lavner & Bradbury, 2010).

The vulnerability-stress-adaptation (VSA) framework (Karney & Bradbury, 1995) integrates all three above-mentioned domains of influence by arguing that changes in relationship satisfaction are governed by the quality of couple interaction, which in turn is a product of the traits and experiences couples bring to the relationship and the stresses that they negotiate.

According to the model, relationship satisfaction is enhanced when couples deal with stressors in constructive ways and worsened when adaptive processes are poor. In addition, the model considers vulnerabilities that each spouse brings into the relationship than can contribute to the experienced stressfulness of the event and affect how couples adapt to new circumstances and situations.

## Consequences of Relationship Satisfaction for General Satisfaction with Life

In their meta-analysis of 93 studies, Proulx, Helms, and Buehler (2007) reveal that relationship satisfaction is a strong predictor of life satisfaction and well-being. Their results indicate that gender, marital duration, the way marital satisfaction and personal well-being are measured, and the year in which the study was conducted moderate the association between marital satisfaction and personal well-being. In his meta-analyses of 26 studies on marital satisfaction and depression, Whisman (2001) found that marital satisfaction was negatively associated with depressive symptoms for both men and women. Studies reveal that low marital satisfaction is not only detrimental for an individual's psychological well-being but also for psychiatric disorders, physical health, and self-esteem.

## Developments in the Study of Relationship Satisfaction

Since the start of research on relationship satisfaction, our understanding of the antecedents and consequences of relationship satisfaction has improved tremendously. Scholars have moved beyond mere snapshots of relationship satisfaction to study patterns of change in relationship satisfaction. Also, they have started to investigate the variability in people's patterns of change in relationship satisfaction, showing the spouses' background and characteristics, life stressors, and transitions while in relationship, and interactional/couple processes comprise the largest domains of influence. Although large progress has been made in understanding relationship satisfaction, there are still many avenues that can be taken to broaden and deepen our understanding of relationship satisfaction. Below I highlight some of them.

### Macro-contexts

Most of the predictors of relationship satisfaction focus on individual or couple characteristics. Of course, the macro-level might also influence relationship satisfaction. This set of factors predicting relationship satisfaction is however the least studied in the field of relationship satisfaction, not in the last part because research designs limit such investigations. Although their potential effects on relationship satisfaction may not be as immediately apparent as some personal factors or interpersonal processes, several contextual factors may influence whether and how couples form their relationship, the obstacles they may be confronted with during their relationship, and the resources they can use to maintain their relationship.

### Dyadic Analyses

Although entering a relationship or getting married constitutes a life transition in which the interwovenness of individual life courses is illustrated most prominently, most research examining relationship satisfaction investigates how individuals function in relationships and how their own evaluations of the relationship are influenced by their own attitudes and behavior. Scholars neglect the fact that one of the defining features of a couple is interdependence – the idea that one's partner's behavior and experiences have the capacity to influence the outcomes of the other partner and vice versa (e.g., Thibaut & Kelley, 1959). An important avenue for future research might therefore be to investigate crossover effects from one partner to assessments of relationship satisfaction of their counterpart.

An individualistic perspective on relationship satisfaction overlooks that women and men within couples share a common set of characteristics with each other that make them more similar to each other than to persons from other couples (Kenny, Kashy & Cook, 2006). Studies suggest that relationship

satisfaction might be more the product of the dyadic environment than individual factors (Johnson & Booth, 1998; Kurdek, 2005). This raises questions concerning the frequently found gender differences in studies that assess relationship satisfaction. A dyadic perspective might reveal that changes in relationship satisfaction occur in tandem and that the reported differences between women and men in prior studies should actually be seen as differences between couples, not between partners within the same couple (e.g., Keizer & Schenk, 2012).

### Long Time Span

Another avenue for future research is to conduct studies that have a long time span.

The majority of studies on relationship satisfaction focus on the early years of the relationship (e.g., Lavner & Bradbury, 2010). These studies are able to detect short-term fluctuations in relationship satisfaction but leave more gradually emerging effects undiscovered. Future research should ideally make use of large time spans to be able to detect such changes in relationship satisfaction.

### Conceptualizing and Measuring Relationship Satisfaction

Many studies that investigate relationship satisfaction make use of a single item to measure relationship satisfaction. Focusing on a single measure of relationship satisfaction may average the complexity of affective changes that occur as time passes and situations and circumstances change. Future research should ideally make use of diverse measures of relationship satisfaction which cover a wider range of dimensions of this concept, such as conflict behavior or separate evaluations of positive and negative aspects (Bradbury et al., 2000; Fincham & Linfeld, 1997).

## Cross-References

- ▶ Adaptation
- ▶ Marital Well-being Measures

## References


Aron, A., Norman, C. C., Aron, E. N., & Lewandowski, G. (2002). Shared participation in self-expanding activities: Positive effects on experienced marital quality. In J. A. Feeney & P. Noller (Eds.), *Understanding marriage: Developments in the study of couple interaction* (pp. 177–194). Cambridge/England: Cambridge University Press.

Biblarz, T. J., & Savci, E. (2010). Lesbian, gay, bisexual, and transgender families. *Journal of Marriage and Family, 72*, 480–497.

Bradbury, T. N., Fincham, F. D., & Beach, S. R. H. (2000). Research on the nature and determinants of marital satisfaction. *Journal of Marriage and the Family, 62*, 964–980.

Bradbury, T. N., & Karney, B. R. (2004). Understanding and altering the longitudinal course of marriage. *Journal of Marriage and Family, 66*, 862–879.

Campos, B., Keltner, D., Beck, J., Gonzaga, G. C., & John, O. P. (2007). Culture and teasing: The relational benefits of reduced desire for positive differentiation. *Personality and Social Psychology Bulletin, 33*, 3–16.

Fincham, F. D., & Beach, S. R. H. (2010). Marriage in the new millennium: A decade in review. *Journal of Marriage and Family, 72*, 630–649.

Fincham, E. D., & Linfeld, K. J. (1997). A new look at marital quality: Can spouses feel positive and negative about their marriage? *Journal of Family Psychology, 11*, 489–502.

Haedey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Towards a dynamic equilibrium model. *Journal of Personality and Social Psychology, 57*, 731–739.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

Huston, T. L., Caughlin, J. P., Houts, R. M., Smith, S. E., & George, L. J. (2001). The connubial crucible: Newlywed years as predictors of marital delight, distress, and divorce. *Journal of Personality and Social Psychology, 80*, 237–252.

Johnson, D. R., & Booth, A. (1998). Marital quality: A product of the dyadic environment or individual factors? *Social Forces, 76*, 883–904.

Karney, B. R., & Bradbury, T. N. (1995). Assessing longitudinal change in marriage: An introduction to the analysis of growth curves. *Journal of Marriage and the Family, 57*, 1091–1108.

Karney, B. R., & Bradbury, T. N. (1997). Neuroticism, marital interaction and the trajectory of marital satisfaction. *Journal of Personality and Social Psychology, 72*, 1075–1092.

Keizer, R., & Schenk, N. (2012) Becoming a parent and relationship satisfaction. A longitudinal dyadic perspective. *Journal of Marriage and Family, 74*, 759–773.

Kenny, D. A., Kashy, D., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford.
Kluwer, E. S. (2010). From partnership to parenthood: A review of marital change across the transition to parenthood. *Journal of Family Theory & Review, 2*, 105–125.
Kurdek, L. A. (1997). The link between facets of neuroticism and dimensions of relationship commitment: Evidence from gay, lesbian, and heterosexual couples. *Journal of Family Psychology, 11*, 503–514.
Kurdek, L. A. (1998). The nature and predictors of the trajectory of change in marital quality for husbands and wives over the first 4 years of marriage for first-married husbands and wives. *Journal of Family Psychology, 12*, 494–510.
Kurdek, L. A. (2002). Predicting the timing of separation and marital quality: An eight-year prospective longitudinal study. *Journal of Marriage and Family, 64*, 163–179.
Kurdek, L. A. (2003). Differences between gay and lesbian cohabiting couples. *Journal of Social and Personal Relationships, 20*, 411–436.
Kurdek, L. A. (2004). Are gay and lesbian cohabiting couples really different from heterosexual married couples? *Journal of Marriage and Family, 66*, 880–900.
Kurdek, L. A. (2005). Gender and marital satisfaction early in marriage: A growth curve approach. *Journal of Marriage and Family, 67*, 68–84.
Lavner, J. A., & Bradbury, T. N. (2010). Patterns of change in marital satisfaction over the newlywed years. *Journal of Marriage and Family, 72*, 1171–1187.
Lucas, T., Parkhill, M. R., Wendorf, C. A., Imamoglu, O., Weisfeld, C. C., Weisfeld, G. E., & Shen, J. (2008). Cultural and evolutionary components of marital satisfaction: A multidimensional assessment of measurement invariance. *Journal of Cross-Cultural Psychology, 39*, 109–123.
Patterson, C. J. (2000). Family relationships of lesbian and gay men. *Journal of Marriage and Family, 62*, 1052–1069.
Proulx, C. M., Helms, H. M., & Buehler, C. (2007). Marital quality and personal well-being: A meta-analysis. *Journal of Marriage and Family, 69*, 576–593.
Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: John Wiley & Sons.
Twenge, J. M., Campbell, W. K., & Foster, C. A. (2003). Parenthood and marital satisfaction: A meta-analytic review. *Journal of Marriage and Family, 65*, 574–583.
Whisman, M. A. (2001). The association between depression and marital dissatisfaction. In S. R. H. Beach (Ed.), *Marital and family processes in depression: A scientific foundation for clinical practice* (pp. 3–24). Washington, DC: American Psychological Association.

## Relationship Transitions

► Marriage, Cohabitation, and Well-Being in 30 Countries

## Relative Deprivation Theory

Michaela Schulze[1] and Rabea Krätschmer-Hahn[2]
[1]Faculty of Arts Sociology, University of Siegen, Siegen, Germany
[2]Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt, Germany

### Synonyms

Downward comparison theory; Paradox of social life; Self-others satisfaction; Social comparison theory

### Definition

Someone is labeled as deprived if he/she is underprivileged in a material or immaterial way. A person will be relatively deprived if he/she feels anger or dissatisfaction because of his/her discrimination in relation to the better situated others. Relative deprivation is, in short, the perceived discrepancy between personal status and the status of some relevant other(s). Without using the concept of quality of life explicitly, the concept of relative deprivation is described from the beginning in terms of quality of life substantially.

### Description

Relative deprivation theory is a widely discussed field of contemporary sociology. A common assumption of this field of research is the fact that the feeling of being disadvantaged is related to a reference group. This feeling will arise from the comparison of oneself to others.

Relative deprivation theory is based on the study "The American Soldier" by Stouffer et al. from 1949. Using interviews, the authors show that dissatisfaction of US soldiers in the 1940s arose from the feeling that other people got unjustly faster promoted than themselves. Besides Stouffer et al., Robert K. Merton, an important American sociologist, also dealt with relative deprivation. He connected the relativeness of the concept deprivation to the concept of reference groups (Merton & Kitt, 1950). The most important question in this respect was which reference group the individual chooses for a reference or comparison. In this respect, Merton identified three types of reference groups: (1) If you want to compare yourself to others, you will choose people with whom you have a relationship. (2) You can choose people who have the same or comparable social status. (3) Finally, you can choose people who will have another (superior or subordinated) social status. Hence, individual's feeling of deprivation is generated by an interpersonal comparison.

Walter G. Runciman built on Merton's views and enlarged the theory of relative deprivation by emphasizing the paradox of social life (Runciman, 1966): People who do not see a chance to possess more than they have will not be dissatisfied with their situation. However, they will be dissatisfied with their situation if they see a chance to improve their situation by comparing themselves with other individuals who seem to be in a better situation. Furthermore, according to Runciman, the degree of relative deprivation can be seen as a result of the type of the deprived feelings. On the one hand, a person can feel individually deprived because he/she compares himself/herself to others. On the other hand, a person can feel deprived collectively because he/she compares himself/herself to his/her peer group or other reference groups.

Based on the definitions above, the following citation points at the relationship between objective and subjective criteria: "People's feelings of resentment, grievance, deprivation or dissatisfaction do not depend, in a direct, simple and obvious way on their objective situation. Sometimes the people who are well-off objectively feel more satisfied with their outcomes than do their poorer cousins. Sometimes the opposite is true" (Crosby, 1984: 53). Empirical research in this field shows that deprivation is any kind of social exclusion that derives from the affiliation to a social marginal group and poverty. Deprivation can contribute to alcoholism, extremism, drug addiction, resignation, depression, suicide, and violence. It is obvious that relative deprivation correlate in some way with absolute (objective) deprivation. From a sociological perspective, the most common application of deprivation theory can be found in studies on poverty (e.g., Atkinson et al., 2002). This combination of poverty research and deprivation theory will be briefly summarized here.

Objective deprivation (absolute deprivation) is used to describe a material disadvantage which can be measured by using several economic or social indicators (e.g., distribution of income). In addition, objective deprivation means a disadvantage concerning personal resources like chances or means. From the perspective of society, they are needed to assure life under certain conditions or a certain position in society. From this perspective, the protection of the socioeconomic status by income, education, and career opportunities is emphasized. The official poverty concept which was adopted by the European Council in 1975 is important here. In this understanding, the poor are defined as "persons, families and groups whose resources (material, cultural and social) are limited so as to exclude them from the minimum acceptable way of life of the member state to which they belong" (Eurostat, 2010, p. 6). As a consequence, a national relative poverty threshold of 60 % of the national median household income was defined at the risk of poverty. However, this concept is still relative in respect to national differences. That is why the World Bank developed an absolute poverty indicator. On this understanding, someone lives in poverty if he/she lives on less than $2 per day.

As described above, relative deprivation might mean that a person feels deprived if he/she compares to other persons and realizes that he/she is disadvantaged (e.g., concerning his/her wishes or expectations). A feeling of

dissatisfaction and disappointment is accompanied by subjective deprivation. In this understanding, the subjectively perceived discrepancy between the personal situation and the situation of others can produce social, political, or economic dissatisfaction of a group or an individual. Karl Polanyi argued that relative differences in economic wealth (understood as subjective deprivation) are more important than absolute differences in wealth, i.e., objective or absolute deprivation (Griffin, 1988: 29). Subjective deprivation is in this respect more important to determine the perceived quality of life, because only a perceived deprived situation causes dissatisfaction. Relative poverty and cross-national differences in living standards are specific forms of relative deprivation from an empirical perspective (Atkinson et al., 2002; Hagenaars, 1986).

Both forms of deprivation result in different social situations. First, a socially unstable situation occurs from the lack of sufficient protection of socioeconomic status. Second, stigmatization is used to describe the lack of a sufficient protection of the social status. Third, the loss of contacts or communication possibilities leads to social isolation. However, critics have argued that relative deprivation theory fails to explain why deprived people fail to take action against their situation and join social movements. On the one hand, it is argued that people tend to act to avoid conflict. On the other hand, it is argued as well that people are aware that these situations may arise with or without social action and that there is no guarantee that an improvement of life conditions will emerge from social action (Kendall, 2011: 560).

The study "The American Soldier" was the starting point for research that emphasizes links between objective disparities in living circumstances of people, the perception of those disparities, and the consequences for social cohesion (Walker & Smith, 2000). Social comparison can be upward or downward. As a consequence, people's discovery of similarities and differences between themselves and others can produce positive or negative affective responses and action (Buunk et al., 1999; Smith, 2000).

## Cross-References

- ▶ Deprivation
- ▶ Multiple Discrepancies Theory
- ▶ Objective and Subjective Deprivation
- ▶ Poverty
- ▶ Subjective Poverty

## References


Atkinson, A. B., Cantillon, B., Marlier, E., & Nolan, B. (2002). *Indicators for social inclusion in the European Union*. Oxford: Oxford University Press.

Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W., & Dakof, G. A. (1999). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology, 34*, 366–375.

Crosby, F. (1984). Relative deprivation in organizational settings. *Research in Organizational Behaviour, 6*, 51–93.

Eurostat. (2010). *Combating poverty and social exclusion. A statistical portrait of the European Union 2010*. Luxembourg: Publication Office of the European Union.

Griffin, D. R. (1988). *Spirituality and society: Postmodern visions*. Albany: State University of New York Press.

Hagenaars, A. (Ed.). (1986). *The perception of poverty*. Amsterdam: North-Holland.

Kendall, D. (2011). *Sociology in our times*. Wadsworth: Cengage Learning.

Merton, R., Kitt, A. (1950). Contributions to the theory of reference group behavior. In R. Merton, & P. Lazarsfeld (Hg.), *Continuities in social research. Studies in the scope and method of "The American Soldier"* (pp. 40–105). Glencoe: The Free Press.

Runciman, W. (1966). *Relative deprivation and social justice: A study of attitudes to social inequality in twentieth-century England*. London: Routledge/Kegan Paul.

Smith, R. H. (2000). Assimilative and contrastive emotional reactions to upward and downward social comparison. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison. Theory and research* (pp. 173–200). New York: Kluwer.

Walker, I., & Smith, H. J. (2000). Fifty years of relative deprivation research. In I. Walker & H. J. Smith (Eds.), *Relative deprivation, specification, development and integration* (pp. 1–12). Cambridge: Cambridge University Press.

# Relative Income and Reference Group Behavior

Ruud Muffels
Faculty of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

## Synonyms

Comparison income; Peer group comparisons; Social comparison group

## Definition

Relative income is the income compared to a particular reference income. The reference income can be the income of oneself but at another time point (habituation) or the income of relevant others at any moment in time (social comparison). The level of subjective well-being and changes therein over time appears affected by absolute as well as relative income levels and changes.

## Description

### Absolute and Relative Income and the Easterlin Paradox

From the rich literature on happiness and subjective well-being, it is known that money does not buy much happiness. This is known as the ▶ Easterlin paradox, which posits that beyond a certain level of GDP per capita, wealthier nations are not better off in terms of subjective well-being than less wealthy nations, whereas within countries, absolute income pays off in terms of happiness though at a marginally diminishing rate (Easterlin, 1974). The paradox is explained by the mechanism of habituation or adaptation (Festinger, 1954; Brickman & Campbell, 1971) according to which people are presumed to be in a sort of "▶ hedonic treadmill": due to rising aspirations, increases in wealth do not lead to similar increases in happiness. The paradox is challenged in a seminal paper by Stevenson and Wolfers (2008) who used the World Values Survey, the Eurobarometer, the General Social Survey, and the Gallup's World Poll Data for various years to examine the relationship between SWB and absolute income. From their findings, they conclude that there is no such paradox, money matters for explaining the substantial differences in the levels of SWB across nations but also within countries, absolute income pays off for happiness. Their analysis and conclusions are questioned again by Easterlin in 2009 (Easterlin & Angelescu, 2009) arguing that the authors mix up the short- and long-term income-happiness relationship. Speculating on the reason for the different results found for the short and long term, the authors refer to the impact of "loss aversion," meaning that ▶ hedonic adaptation is asymmetric, people do not habituate to income losses, whereas they do to income rises but only in the long term.

Another explanation refers to *relative* or *comparison income* and *reference group behavior* because rising incomes are subject to social comparisons with the neighboring Jones's, whose incomes also keep going up. People adapt to their own and their neighbors' increased levels of income by raising their expectations, with the result that no lasting increase in happiness occurs (Headey, Muffels, & Wagner, 2010). Among others, this argument is put forward to explain the Easterlin paradox by the British economist Sir Layard (2005), arguing that people are especially concerned with their relative position, that is, how they fare compared to their peers, instead of how they fare in absolute terms. According to Layard, absolute income only matters at very low levels of income. People with increasing incomes but below their peers' income rises incur a happiness loss reducing the initial gain in happiness associated with the income change confirming the results of other studies (Ball & Chernova, 2007; Ferrer-i-Carbonell, 2005). Many studies use classical survey tools for measurement, but experimental approaches provide new avenues for research on these comparison effects. One such new experimental approach used functional magnetic resonance imaging (fMRI) as

a tool for investigating the effect of relative income on life satisfaction. While being scanned in two adjacent MRI scanners, pairs of subjects had to simultaneously perform a simple estimation task that entailed monetary rewards for correct answers. The research showed that variation in the comparison subject's payment affects blood-oxygenation-level-dependent (BOLD) responses in the ventral stratum. The research showed that activation in the ventral stratum increases in absolute income and – for a given level of absolute income – decreases in lower relative income corroborating the results of more traditional survey research methods (Dohmen, Falk, Fliessbach, Sunde, & Weber, 2011; Fliessbach et al., 2007).

### Income Inequality

At the macro level, the findings of Wilkinson and Pickett on the relationship between income inequality and other measures of country's social and economic performance including subjective well-being (Stiglitz et al., 2009) seem relevant as published in their seminal study "The Spirit Level" (2008). They showed that, using time series and survey data for a large number of countries, more equal societies in terms of a lower ▶ Gini coefficient show better performance in nearly every domain of life including subjective well-being than more unequal societies. The findings were replicated in a recent study using the European Values Survey for 2008 confirming the negative relationship between income inequality measured by the Gini coefficient and subjective well-being (Muffels et al., 2012).

### Reference Group Behavior

There is a long tradition in economics arguing that relative income is more important for people's subjective welfare than absolute income. In the Netherlands, the so-called Leyden School developed a theory of preference formation and ways of direct measurement of preferences (e.g., Kapteyn et al., 1977) arguing that reference group comparison is essential for understanding subjective evaluations of income. Persons compare themselves to others and experience gains in happiness only if their income gains are larger than those of relevant others with whom they compare their own income situation. The main question however is with whom people tend to compare their own situation with. Some studies argued for the comparison with individuals of the same social class (Veenhoven, 1991); persons with the same education (Ferrer-i-Carbonell, 2005), employment (Clark, Frijters, & Shields, 2008), or employer (Brown, Gardner, Oswald, & Qian, 2008); persons of the same age or sex (Firebaugh & Schroeder, 2009); or people who live nearby, including neighbors. Only a few studies use colleagues as a comparison group. Brown et al. (2008) used employer-employee data showing that individuals whose wages rank higher within a firm's wage distribution are more satisfied. Until recently, there was little empirical evidence on the question with whom people in reality tend to compare their situation. Therefore, it is not surprising that most studies have just constructed a reference group from the data by assuming that the reference group consists of people of the same age, sex, income level, occupational status, and/or region (Firebaugh & Schroeder, 2009; Kapteyn et al., 1977). However, a study based on a cross-country comparison of 25 post-transition countries showed that individuals usually compare themselves with colleagues and former schoolmates (Clark & Senik, 2010). The results of these studies generally are that comparison or relative income is negatively related to subjective well-being. The higher the incomes of the comparison or reference group, the lower people's own life satisfaction tends to be. The way relative income and reference group are measured matters for the effect sizes but not for the direction of the effect. Also, Muffels et al. (2012) found significant negative effects of relative income for subjective well-being using the European Values Survey data for 2008.

### Discussion

The issue is still unresolved, what the contribution of absolute versus relative income is to happiness and to what extent people compare themselves with others (reference group) and respond accordingly. Do they compare their

income or their standard of living (status goods) with others, and do they compare with people who are better off or who are worse off? Is the judgment affected by the quality of the relationship with their peers, and do they compare themselves always with the same people or is that dependent on the context in which the comparison is made? A related issue which needs further scrutiny concerns the question whether countries with a higher income inequality are indeed necessarily worse off in terms of happiness as Wilkinson and Pickett (2009) posited (see also Alesinaa, Di Tellab, & MacCulloch, 2004). Also, that evidence suggests that people judge their own welfare level in function of how relevant others, their peers, judge their welfare or income level (preference interdependence). A question not tackled yet is to what extent the impact of comparison income on behavior is different for different groups such as natives and migrants. In most cases, reference groups have been constructed from the data, but idealistically, a better approach would be to measure it using people's answers to questions about the group to whom they compare their income. Including measures in experimental research settings or in large-scale questionnaire would enrich future analyses on reference group behavior not only with respect to its impact on subjective well-being but also for a wide range of economic and social choices. We gave one example of experimental research methods which seem to provide a new avenue for research into the socio-markers of subjective well-being and the way processes of social comparison and adaptation impact on subjective well-being and mental health.

## Cross-References

- ▶ Easterlin Paradox
- ▶ Experimental Design
- ▶ Hedonic Adaptation
- ▶ Inequality
- ▶ Poverty
- ▶ Relative Deprivation Theory
- ▶ Social Comparison

## References


Alesinaa, A., Di Tellab, R., & MacCulloch, R. (2004). Inequality and happiness: Are Europeans and Americans different? *Journal of Public Economics, 88*, 2009–2042.

Ball, R., & Chernova, K. (2007). Absolute income, relative income, and happiness. *Social Indicators Research, 88*, 497–529.

Brickman, P. D., & Campbell, D. T. (1971). *Adaptation Level Theory, ed Appley MH*. Academic, New York, pp 287–302.

Brown, G. D. A., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does wage rank affect employees' well-being? *Industrial Relations, 47*, 355–389.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature, 46*, 95–144.

Clark, A. E., & Senik, C. (2010). Who compares to whom? The anatomy of income comparisons in Europe. *The Economic Journal, 120*, 573–594.

Dohmen, T., Falk, A., Fliessbach, K., Sunde, U., & Weber, B. (2011). Relative versus absolute income, joy of winning, and gender: Brain imaging evidence. *Journal of Public Economics, 95*(4–5), 279–285.

Easterlin, R. A. (1974). Does economic growth improve the human lot? In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramowitz* (pp. 89–125). New York: Academic Press.

Easterlin, R. A., & Angelescu, L. (2009), Happiness and growth the world over: Time series evidence on the happiness-income paradox, IZA Discussion Paper No. 4060.

Ferrer-i-Carbonell, A. (2005). Income and well-being: An empirical analysis of the comparison income effect. *Journal of Public Economics, 89*, 997–1019.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.

Firebaugh, G., & Schroeder, M. B. (2009). Does your neighbor's income affect your happiness? *The American Journal of Sociology, 115*, 805–831.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., et al. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science, 318*(5854), 1305–1308.

Headey, B. W., Muffels, R. J., & Wagner, G. G. (2010). Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness. *Proceedings of the National Academy of Science of the United States (PNAS), 107*, 17922–17926.

Kapteyn, A., Van Praag, B. M. S., & van Herwaarden, F. G. (1977). Twelve thousand welfare functions of income: A comparison of six samples in Belgium and the Netherlands. *European Economic Review, 9*, 283–300.

Layard, R. (2005). *Happiness: Lessons from a new science*. New York: Penguin.

Muffels, R., Dingemans, E., & Skugor, D. (2012). *Money does not buy much happiness – but what have income inequality, modernization and personal values got to do with it*? ReflecT Working Paper Series, 12/001. Tilburg: Tilburg University, (available as SSRN paper), pp. 32.
Stevenson, B., & Wolfers, J. (2008). *Economic growth and subjective well-being: Reassessing the Easterlin Paradox* (IZA Discussion Paper 3654 Institute for the Study of Labor).
Stiglitz, J., Sen, A., & Fitoussi, J.-P. (2009). Report by the commission on the measurement of economic performance and social progress, European Union, Brussels.
Veenhoven, R. (1991). Is happiness relative? *Social Indicators Research, 24*, 1–34.
Wilkinson, R., & Pickett, K. (2009). *The spirit level. Why equality is better for everyone*. London, UK: Penguin Press.

## Relative Income in South Africa

▶ Relative Standing and Subjective Well-Being in South Africa

## Relative Individual Deprivation

▶ Deprivation and Social Exclusion in Europe

## Relative Length of Life Inequality (RLI)

▶ Length of Life Inequality

## Relative Odds

▶ Odds Ratio

## Relative Poverty Rate

Udaya R. Waglé
School of Public Affairs and Administration, Western Michigan University, Kalamazoo, MI, USA

### Synonyms

Poverty incidence; Poverty rate

### Definition

Relative poverty rate indicates the rate or incidence of poverty measured in relative terms. It captures the size of the population in poverty identified using relative poverty thresholds. This size depends on the given distribution of income or other resources used to measure poverty with the applicable poverty threshold determined exclusively from this distribution and those holding or generating resources short of this threshold classified as the "relative poor." Greater inequality in distribution of resources is therefore expected to produce greater relative poverty rates.

### Description

Poverty signifies inadequacy of resources to maintain an acceptable living standard. Operationally, this translates to having income or consumption short of applicable poverty threshold. Poverty thresholds can be absolute or relative, with poverty rates resulting from relative thresholds termed as relative poverty rates. Identifying people (typically families or households) as poor or nonpoor is important for understanding the magnitude of poverty in a specific geographic or administrative unit such as community, region, or country so that appropriate social policies can be formulated to help the poor. To aid in this process, most countries especially in the developing world have established official poverty thresholds.

Ideally, a poverty threshold accurately reflects what it takes to avoid a poor living standard. While what specific goods and services are considered essential is debatable, the standard practice is to create a basket of goods and services including food, housing, and other consumption items following the "basic needs approach" and value its cost at the prevailing market price. Given the centrality of food and housing in this "consumption basket," the process is simplified by adding some allowance based on these food and/or housing costs for other consumption items such as clothing, transportation, medical care, education, and leisure.

Most of these exercises focus on developing absolute poverty thresholds. Poverty thresholds

established in the United States and most other developing countries including Ghana, India, and Mexico fall in this category. Using national poverty thresholds from a sample of underdeveloped countries as a guide, the World Bank has established $1 or $2 a day of income as the international poverty threshold to assess progress in reducing absolute poverty in particular (Ravallion, Dutt, van de Walle, 1991).

The relative poverty threshold, which is used in an increasing number of industrial countries and in comparative research, puts living standard in a relative framework. While this does not deviate from the idea of basic needs or minimally acceptable standard of living, this approach relies on the overall income distribution to determine what is needed to avoid poor living standard. The practice is to use 50 % (or even 60 %) of the median income as the poverty threshold. The assumption is that those with about one half the median income are barely able to maintain this minimally acceptable living standard. Given its connection with distribution of income, greater inequality especially at the lower half of the distribution is likely to result in larger poverty estimates.

### Absolute Versus Relative Poverty?

These absolute and relative poverty thresholds both offer specific merits. At the most fundamental level, an "absolute irreducible core" of resourcefulness is needed to avoid starvation, chronic illness, and other basic human miseries (Sen, 1985). Determining what needs are basic and therefore ought to enter the poverty calculus, however, depends on the specific contexts, time, and physical and cultural requirements (Townsend, 1985). While philosophical debates regarding the appropriateness of absolute and relative poverty thresholds continue, the famous quote from Adam Smith (1776) that poverty is about the ability to "appear in public without shame" suggests that relative needs have been recognized throughout the history of modern economics. Complexities arise, however, since weaker sections of society feel deprived when inequalities are high despite having relatively good economic performance. On the other hand, needs do not attenuate in times of weak economic performance. Further challenges remain as the space used to determine poverty expands to "capabilities" and "social exclusion," focusing on one's "inner" and "societal" resources to function well in society (Sen, 1987; Wagle, 2008). Yet, the argument to relate "absolute capability deprivation" with "relative deprivation" in terms of resources helps reconcile the absolute vs. relative poverty debate.

### Relative Poverty and Quality of Life

Poverty rates have direct connection with quality of life. Regardless of the type of threshold used, the focus is on the ability to maintain an acceptable quality of life, operationally defined as the living standard. More specifically, relative poverty thresholds are determined with the assumption that the adequacy of resources needed to maintain acceptable quality of life depends on the overall distribution of resources in society. Philosophical debates aside, this is reasonable since the normative quality of life depends on what other people are able to maintain especially within a country in which citizens share common economic and political system and resources.

## Cross-References

- ▶ Capabilities
- ▶ Consumption
- ▶ Economic Well-being
- ▶ Income Distribution
- ▶ Low Income
- ▶ Poverty
- ▶ Poverty Lines
- ▶ Poverty Measurement
- ▶ Quality of Life
- ▶ Social Exclusion
- ▶ Social Policy
- ▶ Social Welfare

## References

Ravallion, M., Dutt, G., van de Walle, D. (1991). Quantifying absolute poverty in the developing world. *Review of Income and Wealth, 37*, 345–361.

Sen, A. K. (1985). A sociological approach to measurement of poverty: A reply to Professor Peter Townsend. *Oxford Economic Papers, 37*, 669–676.

Sen, A. K. (1987). *The standard of living*. Cambridge, UK: Cambridge University Press.
Smith, A. (1776). *An inquiry into the nature and causes of wealth of nations* (republished, edited by Edwin Cannan). New York: Modern Library.
Townsend, P. (1985). A sociological approach to measurement of poverty-a rejoinder to Professor Amartya Sen. *Oxford Economic Papers, 37*, 659–668.
Wagle, U. R. (2008). *Multidimensional poverty measurement: Concepts and applications*. New York: Springer.

## Relative Standards Models

▶ Multiple Discrepancies Theory (MDT)
▶ Interpersonal and Intrapersonal Comparisons in Happiness Ratings

## Relative Standing and Subjective Well-Being in South Africa

Dorrit Posel and Daniela Casale
School of Development Studies, University of KwaZulu-Natal, Durban, South Africa

### Synonyms

Comparison income; Relative income in South Africa

### Definition

An individual's relative standing refers to how the individual's economic status compares to the economic status of others. If economic status is measured by income, then an individual's relative standing or relative income can be captured by that individual's rank in the income distribution of the country, for example, while absolute income refers to the individual's level of income. Subjective well-being (SWB) represents how satisfied individuals are with their lives and typically is measured using a Likert scale ranging from very dissatisfied to very satisfied.

### Description

In a seminal study on the determinants of SWB, Easterlin (1974) identified what seemed to be contradictory findings (now referred to as the Easterlin paradox): within countries at any one point in time, an individual's SWB increased as absolute income increased, but across countries over time, SWB did not rise with growing incomes. One explanation for these findings, known as the relative income hypothesis, is that people will not be more satisfied with their lives if their absolute income rises along with the income of others, so that their rank in the income distribution is left unchanged.

Several studies have tested the relative income hypothesis, and many find that both absolute income and relative income are positively correlated with SWB, but the relationship is stronger for relative income. Individuals who rank more highly in the income distribution, or in comparison to a particular reference group, report considerably higher levels of SWB (cf. Clark & Oswald, 1996; Diener, Sandvik, Seidlitz, & Diener, M., 1993; Ferrer-i-Carbonell, 2005; Luttmer, 2005; Michalos, 1985; McBride, 2001).

A key issue that has been investigated in these studies concerns the relevant reference or comparison group for measuring relative standing: when individuals make income comparisons, to whom do they compare themselves? McBride (2001) suggests that both "external" and "internal" reference groups or norms may be relevant. An external comparison group could be people who are similar to the individual in terms of race, age, gender, education, employment, or some combination of these factors. Individuals may also make internal comparisons, for example, with reference to how their current income compares to their income level in the past.

Most studies which explore the relative income hypothesis have used "objective" measures of the individual's relative standing, such as the mean income of the reference group or the individual's rank in the relevant income distribution. However, there are at least two reasons why a subjective measure of the individual's relative standing would provide a better measure with which to test the relative income hypothesis.

First, individuals may not have sufficient information to assess accurately how their income compares to others. For example, individuals may view their income as being lower than that of their reference group and so feel relatively deprived, even if they are actually better off. If relative standing affects individual SWB because of perceptions of relative deprivation or relative advantage, then a subjective measure of relative income – where individuals think they rank relative to others – would be a better measure of relative standing than where individuals actually rank based on their reported incomes. Second, relative income typically is measured using current monthly income (or expenditure). But a ranking based on current income may differ from one based on permanent income, which takes into account the accumulation of income over time. When forming perceptions about their relative standing, individuals are likely to incorporate long-term income flows into their subjective assessments.

Posel and Casale (2011) compare the relationship between objective and subjective measures of relative standing and SWB in South Africa, using data from the 2008 National Income Dynamics Study. This survey was the first nationally representative household survey that collected individual-level information on SWB from all adults in the household, with the further advantage that data on both self-reported income and individuals' perceptions of their relative standing were collected. South Africa is a particularly interesting country to explore the relationship between relative standing and subjective well-being because it has very high levels of inequality. Moreover, following decades of racial discrimination institutionalized during the apartheid era, the distribution of income is heavily skewed in favor of Whites, who represent less than ten percent of the total population. In the years following the first democratic election in 1994, inequality levels have remained very high, but there is evidence that this has been sustained by increasing intra-racial inequality and particularly among Africans (Leibbrandt, Finn, & Woolard, 2012; Leibbrandt, Poswell, Naidoo, & Welch, 2006).

In their study, Posel and Casale (2011) find a wide divergence for South Africa, between where individuals rank in the distribution of (current monthly) income and where individuals think they rank on a six-rung economic ladder for South Africa. Furthermore, Africans who rank in the upper third of the income distribution are considerably less likely than Whites to perceive their economic status as corresponding to the upper two rungs of the economic ladder. One explanation for this race difference is that given the legacy of apartheid, even relatively rich Africans still perceive their economic status as being inferior, particularly when compared to Whites. Another explanation is that people base their assessments in the ladder question on long-term or permanent income, while the income rank is based on (current) monthly income. In the context of large historical inequalities in access to resources, current monthly income may not be a good predictor of long-term income and particularly among Africans.

In testing the relative income hypothesis using regression analysis, Posel and Casale (2011) find that perceived relative standing is much more strongly correlated with SWB than is relative income based on self-reported income and that this is particularly the case among Africans. They find also that an individual's perceived ranking in the village or suburb is more closely related to SWB than is the individual's perceived ranking in the national distribution. This relationship, which is also strongest among Africans, suggests that people may care especially about their status among others who are geographically more proximate. Internal reference norms are also a significant correlate of SWB. However, SWB is more strongly correlated with whether individuals feel that they have moved up the income ranking relative to their family when individuals were age 15, than if they anticipate being better off in the future.

A number of other studies have investigated the correlates of SWB in South Africa, and several have tested the relationship between relative income and SWB using data on self-reported current income (cf. Bookwalter & Dalenberg, 2009; Hinks & Gruen, 2007; Kingdon & Knight, 2006, 2007; Powdthavee, 2007). However, none

of the results in these studies is directly comparable with those in Posel and Casale (2011). Most studies have analyzed data from the 1993 Project for Statistics on Living Standards and Development in which "household-level" information on SWB was captured. In asking respondents "how satisfied is the household with how it lives these days," the question assumes not only that an individual member of the household is able to report objectively on the household's level of satisfaction but, more fundamentally, that there is a unified SWB function at the household level.

## Discussion

The relative income hypothesis predicts that individuals who are advantaged relative to others will report higher levels of life satisfaction or SWB. When measuring relative standing, most studies use objective measures of an individual's relative position, such as the individual's rank in the income distribution. However, these measures may not be a good indication of whether an individual feels relatively advantaged (or relatively deprived). In South Africa, given the legacy of racial inequality, Africans who are relatively well-off based on their current monthly income may still feel relatively deprived particularly in comparison to Whites. Moreover, their current monthly income may indicate a higher relative standing than a ranking based on more long-term income prospects. These factors may explain why among Africans, perceived relative standing has a far stronger relationship with SWB than an objective measure of relative standing based on current income. Further studies are needed to compare objective and subjective measures of relative standing, and their relationship with life satisfaction, in other countries and in different contexts.

## Cross-References

- ▶ Easterlin paradox
- ▶ Income distribution
- ▶ Inequality
- ▶ Norms
- ▶ Likert scale
- ▶ Subjective well-being
- ▶ Survey

## References


Bookwalter, J. T., & Dalenberg, D. R. (2009). Relative to what or whom? The importance of norms and relative standing to well-being in South Africa. *World Development, 38*(3), 345–355.

Clark, A., & Oswald, A. (1996). Satisfaction and comparison income. *Journal of Public Economics, 61*(3), 359–381.

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute? *Social Indicators Research, 28*(3), 195–223.

Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz* (pp. 89–125). New York: Academic.

Ferrer-i-Carbonell, A. (2005). Income and well-being: An empirical analysis pf the comparison income effect. *Journal of Public Economics, 89*, 997–1019.

Hinks, T., & Gruen, C. (2007). What is the structure of South African happiness equations? Evidence from quality-of-life surveys. *Social Indicators Research, 82*, 311–336.

Kingdon, G. G., & Knight, J. (2006). Subjective well-being poverty vs. income poverty and capabilities poverty? *Journal of Development Studies, 42*(7), 1199–1224.

Kingdon, G. G., & Knight, J. (2007). Community, comparisons and subjective well-being in a divided society. *Journal of Economic Behavior and Organization, 64*, 69–90.

Leibbrandt, M., Finn, A., & Woolard, I. (2012). Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa, 29*(1), 19–34.

Leibbrandt, M., Poswell, L., Naidoo, P., & Welch, M. (2006). Measuring recent changes in South African inequality and poverty using 1996 and 2001 census data. In R. Kanbur & H. Bhorat (Eds.), *Poverty and policy in post apartheid South Africa*. Cape Town, South Africa: HSRC Press.

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics, 120*(3), 963–1002.

McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. *Journal of Economic Behavior and Organization, 45*, 251–278.

Michalos, A. C. (1985). Multiple discrepancies theory (MDT). *Social Indicators Research, 16*, 347–413.

Posel, D., & Casale, D. (2011). Relative standing and subjective well-being in South Africa: The role of perceptions, expectations and income mobility. *Social Indicators Research, 104*(2), 195–223.

Powdthavee, N. (2007). Happiness and the standard of living: The case of South Africa. In L. Bruni & P. L. Porta (Eds.), *Handbook on the economics of happiness*. Northampton, MA: Edward Elgar Publishing.

## Relative Subjective Well-Being (RSWB)

▶ Religion and Subjective Well-being in Ghana

## Relatively Poor Countries

▶ Developing Countries

## Reliability

▶ Reliability Generalization

## Reliability Coefficient

Ralph L. Piedmont
Pastoral Counseling Department, Loyola University Maryland, Columbia, MD, USA

### Synonyms

Cronbach's (alpha) coefficient; Internal-consistency reliability; Test-retest reliability coefficient

### Definition

Reliability coefficient is an index of the amount of true variance operating in a set of raw test scores (Aron & Aron, 2003).

### Description

The reliability coefficient is represented by the term $r_{xx}$, the correlation of a test with itself. Reliability coefficients are variance estimates, meaning that the coefficient denotes the amount of true score variance. This is unlike a standard correlation coefficient where, usually, the coefficient needs to be squared in order to obtain a variance (Cohen & Swerdlik, 2005).

To estimate test-score reliability, at a minimum, one needs at least two observations (scores) on the same set of individuals. The correlation between the sets of observations provides a reliability coefficient. From this simple requirement, a wide variety of reliability studies could be designed, such as the test-retest reliability, the internal-consistency reliability (Cronbach, 2004), the split-half reliability (Feldt & Brennan, 1989), the Cronbach's (alpha) reliability (Allen & Yen, 2002), and many more. The above tests are commonly used as a measure of the internal consistency or reliability of a psychometric test score for a sample of examinees, and they are widely used in the social sciences, business, nursing, and other disciplines.

#### Discussion

There are four factors that influence a reliability coefficient. First is the length of the test; the more items a test has, the higher its reliability will be. Because random error is one source of distortion in a test, adding items has the effect of canceling out the influence of this random error. Second is group homogeneity. It is important when assessing reliability that the group being used has as wide a range of variability on the construct as possible. As with any correlation, restriction of range will serve to attenuate covariation (Webb, Shavelson, & Haertel, 2006).

Third is range of ability. Tests are designed to assess constructs within a specific spectrum, known as its range of convenience. It is within this range that test scores will provide the greatest degree of discrimination. However, outside of this range, a test will not yield as high correlations among items. This is because the test items

will be less relevant to measuring the construct. For example, there are different intelligence scales designed to measure people with varying levels of IQ; a test that is useful in assessing IQ with very bright individuals will not be helpful for assessing intelligence with those who are mentally challenged. Similarly, a measure of normal personality would not be useful for assessing Axis II personality disorders.

Finally, method of measurement will also influence the reliability estimate. Different methods tend to generate different estimates of reliability. For example, test-retest methods will generate, usually, lower estimates of reliability than internal-consistency methods (mostly because there is no time interval between assessments); parallel forms will generate lower estimates than when the same form is used because memory effects are removed.

## Cross-References

- ▶ Alpha Reliability
- ▶ Cronbach's Alpha
- ▶ Internal Consistency Reliability
- ▶ Ordinal Alpha
- ▶ Reliability
- ▶ Test-Retest Reliability

## References

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Long Grove, IL: Waveland Press.

Aron, A., & Aron, E. N. (2003). *Statistics for psychology* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Cohen, R. J., & Swerdlik, M. E. (2005). *Psychological testing and assessment: An introduction to tests and measurement* (6th ed.). New York, NY: McGraw-Hill.

Cronbach, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391–418.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (The American council on education 3rd ed., pp. 105–146). New York: Macmillan.

Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2006). Reliability coefficients and generalizability theory. *Handbook of Statistics, 26*, 4–42.

# Reliability Generalization

Eric K. H. Chan
Department of ECPS, University of British Columbia, Vancouver, BC, Canada

## Synonyms

Meta-analysis; Psychometric analysis; Quality of measures; Reliability

## Definition

Reliability generalization (RG) is a meta-analytic procedure for synthesizing reliability estimates and exploring the influence of study characteristics on the variability of the reliability estimates.

## Description

Quality of life measures must possess good psychometric properties. Reliability is a fundamental issue in the development and evaluation of the psychometric properties of quality of life measures. Over the past decade, a large number of reliability studies have been published. Methods such as meta-analysis are needed to synthesize and integrate the published evidence (Cooper, Hedges, & Valentine, 2009). Introduced by Glass (1976), meta-analysis is a statistical procedure for synthesizing research findings. Score reliability may vary across samples and generalizing reliability evidence on the basis of one study may result in overgeneralization. Vache-Hasse (1998) proposed a meta-analytic method called *reliability generalization* (RG) to synthesize score reliability estimates across studies and to identify study characteristics that may influence the variability of the reliability estimates. In RG, researchers statistically aggregate reliability coefficients.

Reliability is defined as the degree to which the observed test scores are consistent. It is

important to note that reliability refers to the test scores but not the test itself. This is reflected in Wilkinson and the American Psychological Association (APA) Task Force on Statistical Inference (1999) that "a test is not reliable or unreliable. Reliability is a property of the scores on a test for a particular population of examinees" (p. 596). Reliability is also an important issue in the interpretation of research findings as reliability has an influence on effect size and statistical power.

Different sources of reliability are available to support the consistency of tests scores. They include internal consistency, test-retest, split half, and parallel forms. Among the different sources of reliability, internal consistency is the most commonly reported source to demonstrate the consistency of test scores in psychological and educational measures (Hogan, Benjamin, & Brezinski, 2000). Similar results have also been found in health-related quality of life measures (Chan, Darmawanti, Mulyana, & Zumbo, 2011).

In the modern, unitary view of validity, different sources of validity evidence should be integrated to support the interpretation and inferences made from test scores (Messick, 1989). Similarly, different sources of evidence should also be collected, accumulated, and integrated to support the reliability of test scores. In other words, instead of simply reporting or collecting evidence from one source of reliability (mostly alpha) and claiming that the measure is reliable, reliability evidence should be gathered and integrated from different sources (not just alpha) and from different populations (e.g., males, females, and individuals with different ethnic backgrounds).

The procedure for conducting an RG study can be briefly summarized in the following five steps: First, begin with a research question (e.g., Reliability Generalization of the Beck Depression Inventory). Second, search for articles on electronic databases (e.g., PsycINFO and MedLINE) and retrieve the articles. Researchers should also create a flowchart outlining the search process (and search terms used), the number of articles retrieved, and the number of articles included in the final synthesis. Third, develop inclusion criteria (e.g., to be included in the RG study, the articles must report reliability estimates) and select articles that meet the criteria. Fourth, develop a coding sheet and code the studies. The coding sheet should include the recording of the reliability coefficients and study characteristics such as publication year, country, sample size, demographic information, and clinical versus nonclinical sample. Ideally, the studies should be coded by at least two independent coders and inter-rater reliability be examined. Fifth, synthesize the reliability estimates and investigate the moderator variables (i.e., the influence of study characteristics on the reliability coefficient).

Publication bias is an issue in meta-analysis. Publication bias occurs when studies are published based on the results (e.g., studies with statistically significant findings are more likely to be accepted for publication in academic journals) rather than their design or methodological quality (Rothstein, Sutton, & Borenstein, 2005). While there is some evidence showing the existence of publication bias in medicine, empirical investigation is needed to see if the issue occurs in psychometric studies. Statistical procedures such as fail-safe N (Orwin, 1983; Rosenthal, 1979) have been developed to examine the issue of publication bias and the procedures have been used by some RG researchers (e.g., Graham & Christiansen, 2009; Graham, Liu, & Jeziorski, 2006). However, more work is needed to investigate the statistical properties of those methods on the examination of publication bias in reliability studies.

RG remains an important area and we expect to see more RG studies published in the academic literature in the years to come. Readers interested in learning more about how to conduct an RG study should refer to the article by Henson and Thompson (2002). Readers interested in the issues on synthesizing reliability coefficients should refer to Feldt and Charter (2006). Additional information on the statistical methods of RG can be found in Mason, Allam, and Brannick (2007); Bonett (2010); and Botella, Suero, and Gambara (2010).

## Cross-References

- ▶ Meta-analysis
- ▶ Psychometric Analysis
- ▶ Reliability
- ▶ Reliability Coefficient

## References


Bonett, D. G. (2010). Varying coefficient meta-analytic methods for alpha reliability. *Psychological Methods, 15*, 368–385.

Botella, J., Suero, M., & Gambara, H. (2010). Psychometric inferences from a meta-analysis of reliability and internal consistency coefficient. *Psychological Methods, 15*, 386–397.

Chan, E. K. H., Darmawanti, I., Mulyana, O. P., & Zumbo, B. D. (2011). Validity evidence and validation practice in papers published in *Value in Health* (1998–2010). Abstract published in *Value in Health*, 14, A151.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage.

Feldt, L. S., & Charter, R. A. (2006). Averaging internal consistency reliability coefficients. *Educational and Psychological Measurement, 66*, 215–227.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3–8.

Graham, J. M., & Christiansen, K. (2009). The reliability of romantic love: A reliability generalization meta-analysis. *Personal Relationships, 16*, 49–66.

Graham, J. M., Liu, Y. J., & Jeziorski, J. L. (2006). The Dyadic Adjustment Scale: A reliability generalization meta-analysis. *Journal of Marriage and Family, 68*, 701–717.

Henson, R., & Thompson, B. (2002). Characterizing measurement error in scores across studies: Some recommendations for conducting "reliability generalization" studies. *Measurement and Evaluation in Counseling and Development, 35*, 113–126.

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement, 60*, 523–531.

Mason, C., Allam, R., & Brannick, M. T. (2007). How to meta-analyze coefficient-of-stability estimates. *Educational and Psychological Measurement, 67*, 765–783.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education/Macmillan.

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157–159.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 118*, 183–192.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustment*. Chichester, West Sussex: Wiley.

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement, 58*, 6–20.

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604.


# Reliability, Statistical


Peggy Schyns
The Netherlands Institute for Social Research, SCP, The Hague, The Netherlands


## Synonyms

Consistency; Parallel tests

## Definition

The reliability of a quality of life measure refers to the degree that the measure yields the same results from measuring the same phenomenon at different points in time.

## Description

To assess the overall quality of a QoL measurement instrument, we look at two main technical criteria: an instrument's validity and reliability. Here we will focus on the latter criterium.

The reliability of a quality of life measure is not something which is immediately apparent by glancing at a specific measure. A general definition of reliability is that it tells us the extent to which the measure yields the same results from measuring the same phenomenon at different points in time. We therefore have to test the reliability of a measurement instrument. The two most well-known techniques are to assess

the ▶ test-retest reliability, or consistency over time, and the ▶ internal consistency of a measure (Punch, 2005).

▶ Test-retest reliability refers to making the same measurement more than once (Babbie, 2004). It means that at least two measurement moments are necessary. For example, if one assesses a person's satisfaction with income at time 0, then the same result should be yielded at time 1. Of course time 1 should be within a range which does not allow real change over time, for example, because of a raise in income. Sometimes the second measurement is carried out at the end of the same questionnaire, sometimes within a month or some months. A longer period of time between two measurements is usually unwise, since real changes in the attitude under scrutiny could have taken place. Larsen et al. (1985) reported test-retest reliabilities ranging between 0.29 and 0.66 for single-item questions and higher reliabilities for multi-item scales. Since it is rather costly and time-consuming to assess the same measurement at different points in time, and to circumvent the danger of measuring real change in the phenomenon instead, researchers usually develop multiple items of the same phenomenon to assess its ▶ Internal consistency.

A measure's ▶ internal consistency concerns the relationship between multiple measures of the same underlying concept. For example, to assess a person's satisfaction with life, Diener developed a multiple-item scale, called the ▶ Satisfaction with Life Scale (SWLS) (Diener et al., 1985). This 5-item scale is developed in such a way that these five items reflect a coherent and consistent picture of a person's satisfaction with life. There are several ways to assess the extent to which a measurement instrument is reliable. Among the most well known in survey or experimental research are parallel tests (where two separate QoL measurement instruments are developed and same measurement results would imply reliability of the underlying concept (Segers, 1987)) and Split-half techniques, of which ▶ Cronbach's alpha is a specific case. For example, for the Satisfaction with Life Scale, a Cronbach's alpha of 0.93 was reported (Larsen et al., 1985).

These types deal with the reliability of the measurement instrument itself, but there are also ways to test the reliability of research workers. For example, also interviewers and coders can make mistakes. To check for this type of unreliability, supervisors can listen in on interviews or can call back a subsample of the respondents to check information (Babbie, 2004). In the case of coding answers from qualitative interviews on QoL, several coders can be assigned to code the same pieces of information. If there is a high agreement between the results, then the inter coder reliability is high.

Finally note that a certain tension exists between a measure's validity and reliability (see Babbie, 2004). To capture a measure in all its nuances is difficult in a complex social world. Take, for example, the concept of quality of life. It is such a broad and rich concept that if we were to design a valid instrument, we could develop hundreds of items which reflect the different dimensions of the concept. However, this may result in losing some of the measure's reliability since each item may not hold under every situational circumstance. Yet, if we want to develop most of all reliable items, some of the richness of the meaning of the concept may get lost, lowering its validity. The two terms are therefore always relative to each other.

## Cross-References

- ▶ Internal Consistency Reliability
- ▶ Qualitative Methods
- ▶ Satisfaction with Life Scale (SWLS), an Overview
- ▶ Test-Retest Reliability

## References

Babbie, E. R. (2004). *The practice of social research*. Belmont, CA: Wadsworth/Thomas Learning.

Diener, E., Emmons, R. A., Larson, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75.

Larsen, R. J., Diener, E., & Emmons, R. A. (1985). An evaluation of subjective well-being measures. *Social Indicators Research, 17*, 1–17.
Punch, K. F. (2005). *Introduction to social research. Quantitative and qualitative approaches*. London: Sage.
Segers, J. (1987). *Methoden voor de sociale wetenschappen*. Assen, The Netherlands: Van Gorcum.

# Reliable Change Index

Martin Guhn[1], Barry Forer[2] and Bruno D. Zumbo[3]
[1]Human Early Learning Partnership, School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada
[2]University of British Columbia, Vancouver, BC, Canada
[3]Department of ECPS & Department of Statistics, Institute of Applied Mathematics, University of British Columbia, Vancouver, BC, Canada

## Synonyms

Critical difference; Clinical significance; Clinically significant change; Significant difference

## Definition

*Reliable Change Index* (RCI) is a concept in measurement and assessment. An RCI is a psychometric criterion used to evaluate whether a change over time of an individual score (i.e., the difference score between two measurements in time) is considered statistically significant. Computationally, RCIs represent a ratio, in which the numerator represents an actual *observed difference score* between two measurements, and the denominator is some form of *standard error of measurement of the difference*. An RCI indicates whether an individual change score (e.g., between a patient's pre-intervention and post-intervention assessment) is statistically significantly greater than a difference that could have occurred due to random measurement error alone.

## Description

The concept of *Reliable Change Index* (RCI) refers to a method that is used to test whether a change over time – that is, the difference score between two assessments of the same person at two points in time – may be considered "reliable" or "(clinically) significant." In particular, RCIs are commonly used to assess whether some condition or construct (e.g., depression; cognitive functioning) changed during an intervention (e.g., between pre-intervention and post-intervention).

The term "reliable change" is used to differentiate change that is reliable in the statistical sense (i.e., change that is statistically significant) from change that may have occurred due to random fluctuation in measurement (e.g.▶ measurement error; Jacobson & Truax, 1991; Maassen, 2004).

Originally, Jacobson and Truax introduced an index to assess change in 1981, based on *Classical Test Theory*, building on previous work by McNemar (1962) and Lord and Novick (1968). The term *Reliable Change Index* was then introduced by Jacobson, Follette, and Revenstorf (1984). Today, the method is commonly referred to as the *Jacobson-Truax Index* or as the *classical approach* to reliable change. However, alternative methods to calculate RCIs have been developed, and the term RCI generally refers to a large number of different variations of reliable change indices that are based on similar concepts.

In essence, an RCI is designed to numerically quantify whether an observed difference between two measurements or assessments may be considered "reliable" or statistically significant. In principle, the formula for an RCI is simple: An RCI is calculated as a ratio, in which the numerator represents the difference between two measurements (e.g., a pre-intervention assessment of depression and a post-intervention assessment of depression) and the denominator represents some form of a *standard error of measurement of the difference*.

$$\text{Reliable Change Index} = x_{(\text{time } 2)} - x_{(\text{time } 1)} / \text{standard error of measurement of the difference} \quad (1)$$

where $x_{(time\ 1)}$ and $x_{(time\ 2)}$ are measurement scores for an individual at different points in time.

In the original formulation of Classical Test Theory, the standard error of measurement of the difference was defined as the following (cf. McNemar, 1962):

$$\text{standard error of measurement of the difference} = \text{variance}_{x1} * (1 - r_{x1}) + \text{variance}_{x2} * (1 - r_{x2}) \quad (2)$$

where $\text{variance}_{x1}$ refers to the variance of all x scores at time 1, $\text{variance}_{x2}$ refers to the variance of all x scores at time 2, $r_{x1}$ is the test reliability at time 1, and $r_{x2}$ is the test reliability at time 2.

The formula shows that the standard error of measurement of the difference takes into account the variance of scores at time 1 and the variance of scores at time 2 as well as the reliability coefficients of time 1 and time 2. All methods for calculating the standard error of measurement of the difference are variations of this formula. The various methods differ in how they substitute – or estimate – the parameters of the formula (e.g., for cases in which the reliability coefficient and/or the variance of scores at time 2 is unknown, some methods assume/propose it is equal to the reliability coefficient and/or variance at time 1).

The size of an RCI is a direct estimate of the statistical significance of the difference score. Given that, in order to obtain an RCI, observed difference scores are divided by the corresponding standard error of measurement of the difference, RCI values are equivalent to standardized ▶ z-scores (i.e., they have a mean of 0 and a standard deviation of 1). Thus, an RCI that is greater than 1.96 denotes a statistically significant difference or, to use the current terminology, reflects a "reliable change." (A z-score of +1.96 corresponds to the 97.5th percentile of a normal distribution; in other words, 95 % of all z-standardized values in a normal distribution are smaller or equal to +/− 1.96. Following the convention of using 5 % as the threshold for statistical significance, an RCI of greater than 1.96 is therefore considered statistically significant).

Despite the simplicity of the generic RCI formula, RCIs have been a matter of debate in the psychometric literature (Hinton-Bayre, 2000; Maassen, 2000; Mellenbergh & Van den Brink, 1998; Temkin, Heaton, Grant, & Dikmen, 1999). Primarily, the debate concerns the question of how to calculate the *standard error of measurement of the difference* – not to be confused with ▶ *standard error* or ▶ *standard error of measurement*, both of which are related, but different concepts. Due to different approaches to conceptualizing and calculating the standard error of measurement of the difference, numerous RCIs exist in the literature. In a review, Perdices (2005) presents eight different approaches to calculating the standard error of measurement of the difference. The different methods for calculating the standard error of measurement of the difference do not reflect trivial differences. In fact, Maassen (2004), Maassen, Bossema, and Brand (2009), and Perdices (2005) demonstrate how applying various RCIs to identical data may lead to a different standard error of measurement of the difference and, consequently, to different conclusions about the "significance" of a change score.

For a user of RCIs, it is critical to explicate the different statistical and conceptual assumptions that underlie the different methods for calculating RCIs. For example, the method proposed by Ley (1972) calculates the standard error of measurement of the difference by inserting the (generic) variance and retest reliability of a test (which, of course, need to be known, e.g., from a test manual/normative sample), as shown in formula (3).

$$\text{standard error of measurement of the difference} = \text{variance}_x * \text{sqrt}(2 * (1 - \text{retest reliability})) \quad (3)$$

In comparison, the formula by Christensen and Mendoza (1986) estimates its parameters from a sample under study, that is, by using the

variances of measurement scores at time 1 and time 2 and the covariance between measurement scores at time 1 and time 2, as illustrated by formula (4).

$$\text{standard error of measurement of the difference} = \text{sqrt}(\text{variance}_{x1} + \text{variance}_{x2} - \text{vairance}_{x1} * \text{variance}_{x2} * \text{correlation}_{x1x2}) \quad (4)$$

Except for in a few special scenarios, these two formulas lead to different results. Furthermore, the two methods are based on different assumptions and are designed for different scenarios. Ley uses a test's generic variance and test-retest reliability (given that they are known) to calculate the standard error of measurement of the difference. This approach assumes that a test's variance and its reliability are equal across samples and across measurement occasions.

Christensen and Mendoza' s approach (1986), on the other hand, takes into account that, for a given sample of interest, the variance of x scores may change between measurement time 1 and time 2, and it considers the covariance between the two measurement time points for the observed sample to be an adequate reference criterion against which to evaluate the significance of an individual change score.

Therefore, Ley's formula seems appropriate if an individual's change over time is to be evaluated in relation to a representative reference sample and if the retest reliability of the measure is known. Mendoza and Christiansen's method, on the other hand, can be used in a scenario in which measurements at time 1 and time 2 for a sample are available and in which individual change scores are to be evaluated in relation to the distribution of difference scores in the given sample. Choosing an appropriate RCI should therefore be guided by conceptual and theoretical considerations. However, as can be seen, in practice, the parameters of a given scenario limit which RCIs can possibly be calculated. As demonstrated above, some RCIs require that the test reliability be known; other RCIs require that the distribution of scores at time 1 and time 2 of a reference sample be known.

## Conclusion and Further Resources

The discussion of the RCI presented here has focused on a generic description of the RCI as a concept in assessment. In addition, we briefly delineated some key challenges and limitations of RCIs that have been discussed in the psychometric literature. Given the substantially different conceptual and computational approaches of the various RCIs for calculating the standard error of measurement of the difference, it is critical that users indicate which RCI they have chosen to use, point out the RCI's potential limitations, and stringently explicate and evaluate the assumptions that underlie the chosen RCI.

## Cross-References

- ▶ Measurement Error
- ▶ Reliability
- ▶ Significance, Statistical
- ▶ Standard Error of Measurement
- ▶ Standard Errors
- ▶ Standard Scores
- ▶ Test-Retest Reliability

## References

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*, 305–308.

Hinton-Bayre, A. (2000). Reliable change formula query. *Journal of the International Neuropsychological Society, 6*, 362–363.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*, 336–352.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Clinical and Consulting Psychology, 59*, 12–19.

Ley, P. (1972). *Quantitative aspects of psychological assessment*. London: Gerald Duckworth.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Maassen, G. H. (2000). Principles of defining reliable change indices. *Journal of Clinical and Experimental*

*Neuropsychology, 22*, 622–632. doi:10.1076/1380-3395(200010)22:5;1-9;FT622.
Maassen, G. H. (2004). The standard error in the Jacobson and Truax reliable change index: The classical approach to the assessment of reliable change. *Journal of the International Neuropsychological Society, 10*, 888–893. doi:10.10170S1355617704106097.
Maassen, G. H. (2005). Reliable change assessment in sports concussion research: A comment on the proposal and reviews of Collie et al. *British Journal of Sports Medicine, 39*, 483–488. doi:10.1136/bjsm.2004.015594.
Maassen, G. H., Bossema, E., & Brand, N. (2009). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology, 31*, 339–352. doi:10.1080/13803390802169059.
McNemar, Q. (1962). *Psychological statistics* (3rd ed.). New York: Wiley.
Mellenbergh, G. J., & Van den Brink, W. P. (1998). The measurement of individual change. *Psychological Methods, 3*, 470–485.
Perdices, M. (2005). How do you know whether your patient is getting better (or worse)? A user's guide. *Brain Impairment, 6*, 219–226. doi:10.1375/brim.2005.6.3.219.
Temkin, N. R., Heaton, R. K., Grant, I., & Dikmen, S. S. (1999). Detecting significant change in neuropsychological test performance: A comparison of four models. *Journal of the International Neuropsychological Society, 5*, 357–369.

## Religion

► Spirituality, Religiosity, and QOL

## Religion and Coping with Child's Death

► Coping with Child's Death Using Spirituality and Religion

## Religion and Politics in the Middle East

► Democracy and Islam in the Middle East

## Religion and Sexism

Lauren E. Maltby[1] and Elizabeth L. Hall[2]
[1]Harbor-UCLA Medical Center, Torrance, CA, USA
[2]Biola University, Rosemead School of Psychology, La Mirada, CA, USA

### Synonyms

Ambivalent sexism; Bigotry and religion; Chauvinism and religion; Gender harassment; Sexist attitudes; Sexist behaviors; Sexist beliefs

### Definition

From a social psychological perspective, sexism includes three different, though related, concepts: sexist attitudes (hostile and benevolent), sexist beliefs (gender stereotypes and gender ideologies), and sexist behaviors (discrimination and its subtype, harassment). In this model, sexist attitudes and beliefs lead to sexist behaviors, or discrimination. Ambivalent sexism theory, introduced by Glick and Fiske in 1996, emphasizes that sexism must be understood within a social context in which men and women are interdependent, that is, they are intimately brought together in the context of heterosexual romance, domestic life, and child rearing. Consequently, rather than being characterized by an attitude of hostility, as is common in most forms of prejudice, sexism is characterized by ambivalence, comprised of both hostility and benevolence. These hostile and benevolent attitudes are closely tied to gender stereotypes and gender ideologies. Gender stereotypes attribute certain characteristics to women and men, while gender ideologies are more general prescriptive beliefs that include stereotypes as well as broader beliefs about proper roles for men and women in society.

The presence of sexist attitudes and beliefs may lead to discrimination. Crosby and Stockdale (2007) state that sex discrimination occurs "when a person is or people are treated unfairly. . . because of gender" (p. 3). It can be seen in unfair levels of

compensation but can also occur in nonmonetary ways, in terms of opportunities, and in interpersonal interactions. Although discrimination is used most often in reference to discriminatory decisions, it actually refers to any kind of treatment that results in people being put at a disadvantage relative to other groups because of their gender. Gender harassment is a subtype of discrimination. Gender harassment consists of behavior and remarks designed to be insulting, degrading, or communicate sexist attitudes about women.

## Description

Although religion has been shown to have many positive effects on the ▶ quality of life for women, sexism has been clearly documented to have negative effects on the quality of life for women. Sexual harassment has been linked to negative physical (e.g., issues of ▶ body image, ▶ eating disturbances, and overall greater physical illness) and psychological (e.g., posttraumatic stress, depression, general clinical symptomology, ▶ life satisfaction, health satisfaction, and psychological distress) outcomes. Religion can contribute to sexism in a number of ways, shaping the gender attitudes, beliefs, and roles of both perpetrators and victims of sexism through gender ideologies, influencing the context in which discrimination occurs, and influencing the experience of sexism. A cross-national study analyzing data from more than 97 countries on gender attitudes, gender well-being, and religion found conclusively that there was indeed a negative relationship between religion and gender ▶ well-being (Seguino, 2011). Interestingly, "in this study, no one religion stands out as consistently more gender inequitable in its effects than all the others ..." (p. 1317). The author goes on to conclude that it is not simply *what* religion one practices but *how* one practices it (i.e., "the intensity of religious belief and the frequency of religious participation" [p. 1317]) that accounts for the negative relationship between religion and gender equality. This confirms earlier findings that individuals high in need for structure demonstrate higher religious fundamentalism and prejudice, including benevolent sexism. With these findings in mind, the research showing a relationship between religiosity and the gender-related attitudes are organized around three broad conclusions: (a) religions often have prescribed gender stereotypes, ideologies, and roles, which are often more conservative than those of the general population; (b) the gender stereotypes and ideologies most strongly correlated with religious variables tend to be benevolent in nature; and (c) the most sophisticated research on religion and sexism suggests that the relationship between sexism and religion is accounted for by specific aspects of religion, rather than by religion understood globally.

### Religion and Patriarchal Attitudes, Beliefs, and Practices

Research from a variety of disciplinary perspectives has documented that religions often play significant roles in structuring gendered beliefs and behavior. It should be noted that research reveals that religions are not monolithic in their approach to gender. In other words, within religious traditions, great variability is often found in beliefs about women's roles, as well as in the leadership roles that are open to women (Gallagher, 2003). Typically, it is the most conservative segments of religions that advocate patriarchal gender roles and exclude women from positions of leadership (Howland, 1999), most often with an emphasis on the "proper" role for women in the home and the control of women's ▶ sexuality.

The variability in gendered beliefs and practices appears to be at least partially accounted for by the cultural practices of the ethnic groups that constitute its practitioners, their degree of acculturation, and trends toward less conservative expressions of religion across time. In contrast to religious expression in their countries of origin, immigrant religions within the United States seem to move toward integration of men and women during worship, acceptance of women in leadership roles (though rarely in the upper hierarchies of leadership), and employment outside of the home. Recent research has documented these trends in Buddhism, Islam, and Hinduism.

Evangelical Christianity has been one of the most studied religious groups in the United States, and this body of research illustrates some of the nuances that must be taken into account in examining the relationship between religiously based beliefs and practices around gender. Research has shown that Evangelical Christians as a whole are more conservative in their views about gender roles than the general population, although there is also considerable within-group variation in these views. Some studies have provided evidence that traditionalist gender ideologies are often more symbolic than actually practiced within evangelical marriages, which tend to function as egalitarian partnerships even while affirming the belief in male headship and female submission (Gallagher & Smith, 1999).

## Religion and Benevolent Sexism

Glick and Fiske's (1996) ambivalent sexism theory is particularly helpful in examining beliefs about women in religious contexts. Religious justifications consistent with benevolent sexism (and in particular the subscale protective paternalism), rather than hostile sexism, are more commonly used to advocate for patriarchal gender roles. In fact, research with both Christian groups in Western nations and Islamic groups in Eastern nations has found that religious variables typically are associated with benevolent or modern sexism and are either not significantly related to hostile sexism, or if they are, the relationship is moderated by gender (Maltby, Hall, Anderson, & Edwards, 2010).

### Sexism and Specific Aspects of Religion

Allport (1954/1979) demonstrated in his analysis of the religion-racism relationship that how a person believes, rather than what he or she believes, is essential. Similarly, Seguino (2011) observed that no one religion is more strongly associated with gender inequitable outcomes than another. Psychologists of religion caution against stereotyping religion as monolithic; religions are multidimensional constructs. In order to draw conclusions about the relationship between religion and sexism, research must include religious variables themselves. In a review of research that has successfully operationalized religion in more nuanced ways, two specific aspects of religion emerge as potentially promising in understanding the relationship between religion and sexism: religious orientation and religious fundamentalism.

*Religious orientation and sexism*. Allport and Ross (1967) first distinguished between intrinsic and extrinsic religiosity in the 1960s. They identified extrinsic religiosity as religiousness with a utilitarian purpose; in this orientation, one uses his or her religious beliefs instrumentally to obtain other ends, such as social connection or solace. In intrinsic religious orientations, however, religion is valued in its own right. Both intrinsically and extrinsically oriented people may behave similarly in terms of their religion (e.g., attending services, reading sacred texts, financial donations); however, the motives and meaning of these behaviors vary greatly depending on the individual.

The importance of individual differences when examining the relationship between sexism and religiosity is exemplified in a study by McFarland (1989) of American undergraduate students. McFarland (1989) found that, contrary to expectations, intrinsic religiosity was negatively correlated to sexist attitudes toward women, and extrinsic religiosity was positively correlated. Burn and Busso (2005) also found that once intrinsic religiosity was statistically controlled, scriptural literalism ceased to be a significant predictor of benevolent sexism. Intrinsic religiosity and scriptural literalism did account for unique portions of variance in the protective paternalism subscale, but not in either of the other two benevolent sexism subscales. These results suggest that religious beliefs (such as scriptural literalism in this case) may serve to reinforce the protective paternalism component of benevolent sexism but also suggest that individual differences (such as intrinsic religiosity) mediate the relationship between sexism and religious beliefs.

*Religious fundamentalism*. Another specific aspect that has emerged as promising in the study of religion and sexism is that of religious fundamentalism. Fundamentalism is a

cognitively rigid style of holding beliefs. Altemeyer and Hunsberger (1992) defined religious fundamentalism as the belief that "there is one set of religious teachings that clearly contains the fundamental, basic, intrinsic, essential, inerrant truth about humanity and deity; that this essential truth is fundamentally opposed by forces of evil which must be vigorously fought; that this truth must be followed today according to the fundamental, unchangeable practices of the past; and that those who believe and follow these fundamental teachings have a special relationship with the deity" (p. 118). McFarland (1989) found that when he controlled for fundamentalism, the relationship between both intrinsic and extrinsic religiosity and sexism disappeared for women (although not for men).

Although the body of literature exploring the role of fundamentalism in the religion-sexism relationship is still relatively small, much insight can be gleaned from the larger body of research on the religion-prejudice relationship. Hunsberger (1995) demonstrated that both fundamentalism and quest orientation are more significant in predicting prejudice than religious beliefs. "In the end, it would seem that it is not religion per se, but the ways in which individuals hold their religious beliefs, which are associated with prejudice" (p. 113). Laythe, Finkel, and Kirkpatrick (2002) found that once authoritarianism was statistically controlled, fundamentalism became a significant predictor for some forms of prejudice (but not others). They concluded that authoritarianism, which has been shown to be reliably related to prejudice, has two components, namely, fundamentalism and Christian belief content, and that these two components are differentially related to prejudice. These findings in the context of ▶ prejudice and religion suggest that the relationship between religion and sexism may likewise be multifaceted.

*Other factors*. Although religious orientation and fundamentalism have clearly emerged in the literature as the most promising aspects of religion to consider in understanding the relationship between religiosity and sexism, a host of other, more general factors have also emerged as influential in this relationship. Several studies have demonstrated that the relationship between religion and sexism is moderated by gender (Maltby et al., 2010; McFarland, 1989). With the exception of McFarland, studies have found that (a) religious men hold more traditional beliefs about gender and gender roles than religious women and (b) the content of religious beliefs seems to be related to sexism for men, but not for women (Maltby et al.). Women's higher participation in religious practices may be the vehicle transforming the relationship between religious beliefs and sexism, but more careful research is necessary to confirm this hypothesis.

In addition to gender, acculturation, ethnicity, cultural and community ▶ norms, and geographic location/city size, all affect the way religion functions to influence sexism. For instance, Taiwanese immigrant women have been shown in some cases to use Buddhism and Christianity, both religions associated with traditional attitudes toward women, to carve out independent selves from their families. Finally, it is important to note that religious communities are not immune to the effects of time. Studies have shown that within Christianity, attitudes about women's roles have converged between denominations and that attitudes toward women's roles had become less traditional over time for Jews and Christians alike. More longitudinal studies are needed to understand cohort and historical effects on the relationship between religion and sexism.

## References

Allport, G. W. (1954/1979). *The nature of prejudice*. Cambridge, MA: Perseus Books.

Allport, G. W., & Ross, J. M. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology, 5*, 447–457.

Altemeyer, B., & Hunsberger, B. (1992). Authoritarianism, religious fundamentalism, quest, and prejudice. *Journal for the Scientific Study of Religion, 35*, 56–64.

Burn, S. M. & Busso, J. (2005). Ambivalent sexism, scriptural literalism, and religiosity. *Psychology of Women Quarterly, 29*(4), 412–418.

Crosby, F. J., & Stockdale, M. S. (2007). Introduction: understanding sex discrimination. In F. J. Crosby, M. S. Stockdale, & S. A. Ropp (Eds.), *Sex discrimination in the workplace* (pp. 3–5). Malden, MA: Blackwell Publishing.

Gallagher, S. K. (2003). *Evangelical identity and gendered family life*. New Brunswick, NJ: Rutgers University Press.
Gallagher, S. K., & Smith, C. (1999). Symbolic traditionalism and pragmatic egalitarianism: Contemporary evangelicals, families, and gender. *Gender and Society, 13*, 211–233.
Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*, 491–512.
Howland, C. W. (Ed.). (1999). *Religious fundamentalisms and the human rights of women*. New York: Palgrave.
Hunsberger, B. (1995). Religion and prejudice: The role of religious fundamentalism, quest, and right-wing authoritarianism. *Journal of Social Issues, 51*, 113–129.
Laythe, B., Finkel, D., & Kirkpatrick, L. A. (2002). Predicting prejudice from religious fundamentalism and right-wing authoritarianism: A multiple-regression approach. *Journal for the Scientific Study of Religion, 40*, 1–10.
Maltby, L. E., Hall, M. E. L., Anderson, T. L., & Edwards, K. (2010). Religion and sexism: The moderating role of participant gender. *Sex Roles, 62*, 615–622.
McFarland, S. G. (1989). Religious orientation and the targets of discrimination. *Journal for the Scientific Study of Religion, 28*, 324–336.
Seguino, S. (2011). Help or hindrance? Religion's impact on gender inequality in attitudes and outcomes. *World Development, 39*, 1308–1321.

## Religion and Subjective Well-being in Ghana

Isaac Addai[1] and Jelena Pokimica[2]
[1]Social Science & Humanities Department, Lansing Community College, Lansing, MI, USA
[2]Case Western Reserve University, Cleveland, OH, USA

### Synonyms

Absolute subjective well-being (ASWB); Relative subjective well-being (RSWB)

### Definition

In defining well-being, social scientists have dedicated a significant amount of scholarly energy over the past three decades; however, what constitutes well-being and how to measure it continues to be the source of considerable debate (Camfield, 2004). The general definition of well-being is captured by McAllister (2005, p. 2) as:

> More than the absence of illness or pathology [. . .with] subjective (self-assessed) and objective (ascribed) dimensions. It can be measured at the level of individuals or society [and] it accounts for elements of life satisfaction that cannot be defined, explained or primarily influenced by economic growth.

Generally, three major approaches to well-being can be discerned from the literature. The first approach focuses on basic and universal needs that are supposed to meet individuals' life needs for survival (Baumeister & Leary, 1995). The second approach projects well-being in terms of people's hedonic experiences (Kahneman, 1999). And the third assesses well-being in terms of individuals' perceptions of their own life referred to as ▶ subjective well-being (SWB). In this approach, participants are asked to make a global evaluation of their own lives (Andrews & Whithey, 1976).

Although definitions vary, SWB is portrayed as a measure that attempts to capture the overall sense of well-being, including ▶ happiness and satisfaction of life as a whole and other domains of life. It is nonobjective and often broad in its definition. In the context of Sub-Saharan Africa, subjective well-being is defined by one's perceived living conditions, either absolute or relative. Absolute subjective well-being (ASWB) is characterized by one's own perceptions of living conditions (how good or bad they are), and relative subjective well-being (RSWB) is characterized by perceptions of one's own living conditions compared to those of others (better or worse). People's outlook on living conditions is often shaped by sociocultural factors that figure in the provision of social services.

### Description

The upswing in scholarly work on ▶ subjective well-being is heightened by a corpus of empirical

work in three interrelated areas, namely, (a) empirical confirmation of the salience of the subjective view in evaluating ▶ quality of life (Andrews & Withey, 1976), (b) the realization that well-being transcends economic prosperity (Putnam, 2000), and (c) scientific evidence that individuals' sense of their own well-being is a valid and reliable measure of quality of life (Diener, Eunkook, Lucas, & Smith, 1999).

The last four decades have witnessed studies on understanding predictors of SWB (Dolan, Peasgood, & White, 2008). One factor which has consistently been implicated in how people gauge their own well-being is religion (Lim & Putnam, 2010). The dominant finding from the extant literature is that religious people tend to report better well-being. Although the mechanism through which religion associates with well-being continues to be debatable, two perspectives can be deduced from the literature. One strand of investigations imputes the superior well-being outcomes of religious people to social support which individuals derive from their membership in religious organizations (Krause, 2008), while others assign this credit to private and psychological aspects of religion (Greeley & Hout, 2006).

Though religion is reported to be the bedrock of African societies (Opoku, 1978), studies examining the association between religion and well-being in Africa are lacking in the literature. We conducted a 2008 Afrobarometer-based study (N = 1,200) to examine associations between religion and SWB in Ghana, mainly, if any significant differences among religious affiliation groups and religious importance groups on SWB existed, and if yes – how did groups vary?

We looked at absolute and relative SWB outcomes and six major religious groups: none/traditional, other Christian (composed of Christians only, i.e., respondents saying only "Christian" in the survey without identifying with a specific subgroup; Orthodox; and Coptic), Catholic, Protestant, Muslim, and Evangelical/Pentecostal. Importance of religion was also considered: "not at all or not very important," "somewhat important," and "very important."

As per Dickson (1975) and Ewusi (1976), because (a) ethnicity represents people's cultural identity and informs variations in their ethnic group positioning (Akan, Ewe/Anglo, Ga/Adangbe, and other smaller Ghanaian groups), (b) educational attainment forms part of the socioeconomic background and captures economic disparities in the country (no education, elementary, high school, and bachelor degree or higher), and (c) sharp differences in regional development exist (Western/Central, Greater Accra, Volta, Eastern/Ashanti/Brong Ahafo, and Northern/Upper East and West), we consider those factors as important in Ghanaians' perceptions of SWB.

Significant religious affiliation and important group differences on both types of SWB are found to be present. The relevance of religious affiliation in Ghanaians' perceptions of living conditions, therefore, tends to persist whether we focus on ASWB or RSWB. As a result, these findings lend credence to a notion that religion significantly molds people's SWB in Ghana.

While religion seems to induce an unfavorable general outlook on living conditions among Ghanaians affiliated with the none/traditional religion, it tends to suggest a favorable sense of living conditions among Ghanaians affiliated with Muslim, other Christian, and Evangelical/Pentecostal religions, especially in terms of ASWB. Similar picture is painted for RSWB, with the exception that none/traditional and Muslim groups did not significantly differ but the none/traditional and Protestant ones did (the first group experienced a less favorable RSWB).

Factors sustaining differentiations in SWB among the religious groups may not be easy to discern. Possible reasons may be differences in socioeconomic characteristics, such as the level of education or residential patterns that characterize these groups. Differences in teachings, practices, and ▶ norms among the groups may be other possible explanations. For instance, while the none/traditional religious group tends to espouse conservative gender roles and discourage formal education, especially among women, the opposite is the case among Evangelicals/Pentecostals and other Christians.

About one fifth of adherents to no religion or African traditional religions live in the Northern and Upper regions of Ghana where ▶ poverty levels are rather high compared to those in the Southern region. The North is also largely populated by Muslim communities, though some of them live in the South where they tend to be found in some of the poorest parts of Ghana's urban areas. These educational and residential pattern differences could figure in some of the less favorable SWB experiences of some groups, such as the none/traditional group, relative to the other ones. Theoretically speaking, teachings of various Christian denominations in Ghana, with their emphasis on salvation and prosperity, provide their members with a sense of hope even when their real life situations may be quite challenging. The notion of suffering on the Earth (postponement of life's riches), as some Christian groups (e.g., Evangelical and Pentecostal) recount for future salvation, may also provide some justification for the perceived interpretation of their living situations as better than those of their non-Christian counterparts (e.g., none/traditional).

The other Christian religious group garnered most positive perceptions of both ASWB and RSWB. Most of the other Christian respondents are from local Ghanaian Christian church groups tending to blend some of the traditional religious practices with the Protestant ones. Their day-to-day outlook on life events and well-being tends to be seen more in distinctive religious terms. In Ghana, these groups espouse more inward-oriented and exclusive culture or what Putnam (2000) calls "bonding social capital." These groups tend to have strong congregational ties and a tight network.

The positive outlook on SWB among other Christians may be further explained in terms of how their members interact. They are inclined to offering social and economic support to members in their community at times of need. Some of these groups have their own farms and healing camps and tend to be isolated from the general society, depending mostly on each other. The sense of support from the community coupled with members' willingness to help the less fortunate may be the rationale behind the observed positive outlook on SWB among other Christians.

Groups in which religion is or is not important also differ significantly. This is mainly the case for the group that does not consider religion very or at all important and the group that considers it very important (two groups on extreme ends) with the latter experiencing a more desirable SWB (of either type). The salience people attach to religion seems to give them some psychological comfort in their sense of living conditions.

## Discussion

Considering the documented link between living conditions figuring in people's SWB and stability of a nation (Böhnke, 2004), the impact of religious affiliation and importance on SWB in Ghana is an important finding for the nation. In the short term, religion can be a vital asset which policy makers can utilize for the stability of the country. In the longer term, there is a need to understand whether teachings and practices, norms, and beliefs of the various groups are a source of comfort and inspiration to members or a hindrance, as far as their experiences of SWB are concerned. This demands a continuous inquiry.

Sociodemographic and regional factors further significantly impact people's SWB in Ghana. Ewe/Anglos seem to have less favorable SWB experiences than Akans and Ga/Adangbes, for example. The more developed South is dominated by Akans having better life opportunities. Many Ga members live in Greater Accra, while Ewes live in the less developed Volta region. The difference in well-being among these ethno-regional groups is further supported by a distribution of social amenities such as schools, hospitals, and the like. Using a distribution of facilities in the country along a continuum of the most to the least developed, Dickson (1975) puts forth the following region list: (1) Greater Accra, (2) Ashanti, (3) Eastern, (4) Central, (5) Western, (6) Volta,

(7) Brong/Ahafo, (8) Northern, and (9) Upper East and West.

Bridging development gaps in social and economic realms (educational, regional) in many parts of the country where different religious groups live would give the majority of Ghanaians better life chances and more encouraging SWB outlooks. Achieving such noble goals, under the umbrella of people's ultimate quality of life, would demand a carefully coordinated effort between the private sector, civil society, and the government, and possible decentralization efforts. For many at the microlevel left out of the benefits from the economic growth and political stability that has been happening at the macrolevel in Ghana since the 1980s, district assemblies tend to be their last hope in gaining from the transformations in the country. These assemblies can work with religious organizations to champion educational and community development projects which could go a long way in providing an optimistic outlook on living conditions among the Ghanaians.

## References


Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being*. New York: Plenum.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497–529.

Böhnke, P. (2004). *Perceptions of social integration and exclusion in an enlarged Europe*. Luxembourg, Europe: European Foundation for the Improvement of Living and Working Conditions. Retrieved from http://www.eurofound.europa.eu/publications/htmlfiles/ef03106.htm.

Camfield, L. (2004). Measuring SWB in developing countries. In W. Glatzer, S. Von Below, & M. Stoffregen (Eds.), *Challenges for the quality of life in contemporary societies*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Dickson, K. B. (1975). Development planning and national integration in Ghana. In D. R. Smock & K. Bentsi-Enchill (Eds.), *The search for national integration in Africa*. NY: Free Press.

Diener, E., Eunkook, M. S., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology, 29*, 94–122.

Ewusi, K. (1976). Disparities in levels of regional development in Ghana. *Social Indicators Research, 3*, 2–5.

Greeley, A., & Hout, M. (2006). *Happiness and lifestyle among conservative Christians. The truth about conservative Christians* (pp. 150–61). Chicago: University of Chicago Press.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York: Russell Sage.

Krause, N. (2008). *Aging in the church: How social relationships affect health*. West Conshohocken, PA: Templeton Foundation Press.

Lim, C., & Putnam, R. D. (2010). Religion, social networks, and life satisfaction. *American Sociological Review, 75*, 914–933.

McAllister, F. (2005). Well-being concepts and challenges: Discussion Paper, SDRN.

Opoku, K. A. (1978). *West African traditional religion*. Accra, Ghana: FEP International Private Ltd.

Putnam, R. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon and Schuster.


## Religion in Adolescents

► Adolescents and Spirituality



## Religion, Psychological Well-Being, and Health

Marta Elliott[1] and Michael J. Doane[2]
[1]University of Nevada, Reno, NV, USA
[2]Interdisciplinary Ph.D. Program in Social Psychology, University of Nevada, Reno, NV, USA

### Synonyms

Religiosity and well-being; Well-being, subjective

## Definition

Religiosity refers to a broad set of beliefs and behaviors centering on the sacred.

## Description

### Introduction

Over the past several decades, there has been an increasing interest among social scientists in documenting the relationship between religion and health, including both psychological and physical health outcomes. The purpose of this entry is to provide an overview of the field with particular attention to theory and evidence regarding specific mechanisms that explain why religion tends to be associated with better mental and physical health. The bulk of this entry will be based on research conducted in the United States but will also include a small body of cross-national studies that compare the nature of the religion/health association between countries. Although the list of potential mediators is extensive, this entry will focus on religion as a source of social identity, of ▶ social support, and of ▶ attitudes, beliefs, and behaviors that promote good health as well as aid individuals in coping with challenging life circumstances. In addition, this entry will note studies that have found religion to be harmful, will outline the weaknesses apparent in the religion/health literature, and will suggest directions for future research.

*Social Identity*: Individuals derive social identities from groups to which they belong that have particular meaning or significance to them, and a positive social identity is associated with greater ▶ psychological well-being (Haslam, Jetten, Postmes, & Haslam, 2009). Although one in five American adults have no religious affiliation (Pew Research Center, 2012), those who do may access a positive social identity associated with their group membership that offers a powerful ▶ sense of belonging. Strength of religious identity partially mediates the association between religious participation and ▶ subjective well-being (Greenfield & Marks, 2007), and the capacity for religious identity to improve well-being increases as a function of the fit between the individual and his or her religious congregation's characteristics (Hayward & Elliott, 2009). It should not be forgotten, however, that overidentification with one's religious group may foster ▶ prejudice and discrimination toward religious out-groups (Mael & Ashforth, 2001), and those who identify with certain religions sometimes suffer persecution in the form of confinement, exile, and even death (Grim & Finke, 2007).

*Religious Motivations*: Motivations for being religious are classified as either intrinsic or extrinsic, and theses religious motivations have been linked to psychological well-being. Those who are intrinsically motivated adhere more closely to religious doctrine, integrate religion into their daily lives, and seek personal meaning in religious activities, whereas the extrinsically motivated use religion for nonreligious purposes, such as seeking security in the face of personal problems or finding friends who elevate one's sense of support and/or social status. Intrinsic religious orientation is associated with less frequent perceptions of ▶ stress, less ▶ anxiety, and depression (Masters & Bergin, 1992), whereas extrinsic religious orientation is associated with greater self-doubt, anxiety, and depression (Bergin, Masters, & Richards, 1987). However, people who endorse an extrinsic orientation toward religion may have worse well-being not because they use religion for non-religion purposes per se but because they lack social support in their everyday lives.

*Social Support*: ▶ Social integration and support are long established predictors of good mental and physical health (Cohen, 2004). As members of religious organizations, individuals may enjoy a sense of belonging as well as perceive and receive social support from religious leaders and fellow congregants. Participation in religious activities embeds the individual within a ▶ social network of like-minded others, an environment that may foster especially supportive relationships that improve well-being (George, Ellison, & Larson, 2002; Krause, Ellison, & Wulff, 1998). In fact, the benefits of religious social support are independent of the

positive influence of general social support on health and mediate the association between religious service attendance and well-being (Doane, 2013). However, there is some evidence that negative social interactions within religious settings increase ▶ distress among clergy and elderly Presbyterians (Krause et al., 1998).

*Positive Emotions*: Positive emotions have been linked to good mental (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009) and physical health (Diener & Chan, 2011). Religious activities, including prayer and service attendance, may produce positive emotions including awe, forgiveness, gratitude, hope, and joy (Levin, 2010). Developing research in this area includes efforts to distinguish between ▶ positive emotions fostered within religious settings (e.g., awe and forgiveness) from those that are likely unrelated to religion (e.g., amusement), as well as delineating which specific aspects of religion cultivate positive emotional experiences. Nonetheless, religious engagement has also been linked to negative emotions, such as anxiety, guilt, self-doubt, and shame (Koenig, George, Blazer, & Pritchett, 1993).

*Health Behaviors*: Religiosity may benefit physical well-being by promoting a healthy lifestyle (e.g., treating the body as a temple). For example, many religions discourage the use of alcohol and drugs (Levin, 2010) while encourage dietary restrictions, regular ▶ exercise, and the utilization of ▶ health care (Chatters, 2000). However, there is some evidence that individuals who subscribe to fundamentalist beliefs have worse physical health than their more liberal counterparts do (Green & Elliott, 2010) perhaps because they defer personal control over their health to God, whom they perceive to in control of their daily lives.

*Sense of Divine Control*: A substantial body of research has established mental and physical health benefits of possessing the belief that one is able to determine important outcomes in life (Gadalla, 2009), as well as health risks of attributing control to external sources, such as luck or fate (Gale, Batty, & Deary, 2008). The influence on health of the belief that God has ultimate control (i.e., "divine control") is more complex, in part because it suggests that individuals who believe in divine control have lower personal control and therefore may not engage in active coping when it might otherwise help them (Ellison, 1991). However, Schieman (2008) found no association between senses of divine versus personal control among deeply religious individuals. In fact, belief in divine control may enhance one's sense of personal control (Pargament, 1997) perhaps because it is perceived as God's gift of free will. At the very least, belief in divine control may offer comfort and hope in an otherwise seemingly uncontrollable world. However, the health effects of belief in divine control are likely modified by beliefs about what God intends when exerting control. For instance, the belief that God causes stressful events in one's life as a form of punishment is associated with greater depression and lower ▶ life satisfaction (Bjorck & Thurman, 2007).

*Sense of Coherence*: ▶ Sense of coherence refers to the perception that the world makes sense, is manageable, and has meaning that assists people in coping with stressors, having favorable association with mental and physical health (Eriksson & Lindström, 2006). Religion can be a major source of sense of coherence (George et al., 2002), though it has not been directly tested as a mediator between religiosity and well-being. However, Ellison (1991) found that a closely related concept, existential certainty, mediated the relationship between service attendance and well-being, as well as buffered the negative influence of traumatic life events. In addition, Bennett and Elliott (2011) found that a form of prayer assists with finding meaning from traumatic experiences. They compared written narratives addressed to God, a friend, or to no one in particular, wherein the "prayer" texts contained more causal and insight words used to described traumatic life experiences. Furthermore, research on the lack of existential certainty (i.e., religious doubt) has been linked to increased somatic symptoms of depression and decreased satisfaction with health (Krause & Wulff, 2004). However, it may also be that an excess of existential certainty characterized by rigid and inflexible religious beliefs could interfere with

individuals' capacities to deal with challenging circumstances. In fact, individuals with more orthodox religious beliefs tend to be more distressed perhaps because when confronted by stressors that test their strict belief systems, they feel threatened and engage in maladaptive coping (Dezutter, Soenens, & Hutsebaut, 2006).

*Religious Coping*: Several of the mediators described thus far potentially aid individuals in coping with stressors in their lives, including religious social identities, religious social support, sense of divine control, and sense of coherence. In addition, Pargament, Smith, Koenig, and Perez (1998) have identified specific positive and negative coping patterns associated with religion. Positive religious coping reflects a secure relationship with religion, including seeking support from religious members and using religion to shift focus away from a stressor, and predicts less psychological distress (Pargament et al., 1994). In contrast, negative religious coping refers to a less secure relationship with religion, including confusion or dissatisfaction with a higher power and negative interactions with religious members, and predicts higher levels of distress (Pargament et al.).

*Cross-National Research on Religion and Well-Being*: Although most research on religion and well-being is based on single-nation studies and most of these studies are from western industrialized countries, particularly the USA, there are a handful of studies that examine the association between religion and well-being in a cross-national perspective. These studies utilize multilevel frameworks to assess how religiosity influences well-being as a function of country-level characteristics. For example, Elliott and Hayward (2009) found that in countries where religious expression is highly regulated, public expression of religiosity (i.e., attendance at religious services) reduces life satisfaction, whereas one's private religious identity is more strongly predictive of greater life satisfaction than is the case in countries where religious expression is relatively free. Several other studies have shown that religion is more beneficial for well-being when religious beliefs or religious affiliations are widely shared. For example, high individual religiosity is more strongly associated with benefits for subjective well-being when societal level of religiosity is also high (Diener, Tay, & Myers, 2011). In addition, the positive association between religiosity and physical health declines as secularism increases (Nicholson, Rose, & Bobak, 2009). Future research on religion and well-being in a cross-national perspective should continue to incorporate between-country differences in key aspects of social life that relate to religious experiences and quality of life.

## Limitations

The primary weakness of religion and health research is its typical reliance on cross-sectional data to test hypotheses that presume religiosity precedes good health in time. Although the theories undergirding the causal mechanisms outlined above make a strong case for religion affecting health, there are also convincing arguments and evidence in support of the idea that mental and physical health predicts religious beliefs and behaviors. For instance, people with serious medical conditions turn to religion for support and comfort (Koenig, Larson, & Larson, 2001) although poor health may preclude active religious involvement. This paradox might explain why people diagnosed with ▶ mental illness are less likely to have a religious affiliation or attend services but more likely to engage in private prayer (Lindenthal, Myers, Pepper, & Stern, 1970). A second limitation of the religion and health research is the tendency for researchers to focus on one or two mechanisms linking religion to health with few if any studies assessing a comprehensive model including the many mechanisms outlined in this essay. Finally, too much research on religion and health is independent of mainstream research on social factors and health, such as studies that consider the health effects of experiences of discrimination, conditions on the job, ▶ interpersonal relationships, and neighborhood characteristics. Research on social factors and health would benefit from more frequently integrating religiosity into their theoretical and empirical models. No doubt many fascinating interrelationships exist between various types of stressors and religious resources that have yet to be explored.

## Conclusion

Although research on religion and health is relatively new to the scene of medical sociology and health psychology, a great deal of evidence has mounted in a rather short time about the mostly salutary effects of religion on human well-being. Future research on religion and health would benefit from the collection of longitudinal panel survey data that includes measures of multiple dimensions of religiosity as well as various types of mental and physical health outcomes, including assessments of well-being and good health as well as serious mental and physical illness. Even better, studies should include multiple types of stressors that have the potential to affect health, including traumatic life events and chronic strains, to test how religiosity may buffer, or exacerbate, their effects on health. With such data, (a) issues of causal direction between religion and health may be addressed, (b) the potentially complex interrelationships between mechanisms linking religion and health may be explored, and (c) research on religion and health may be integrated into the broader fields of social science research that concern themselves with social factors and health.

## Cross-References

► Life Satisfaction, Concept of

## References

Bennett, P., & Elliott, M. (2011). God give me strength: Exploring prayer as self-disclosure. *Journal of Religion and Health*. doi:10.1007/s10943-011-9460-1.

Bergin, A. E., Masters, K. S., & Richards, P. (1987). Religiousness and mental health reconsidered: A study of an intrinsically religious sample. *Journal of Counseling Psychology, 34*(2), 197–204.

Bjorck, J. P., & Thurman, J. W. (2007). Negative life events, patterns of positive and negative religious coping, and psychological functioning. *Journal for the Scientific Study of Religion, 46*(2), 159–167.

Chatters, L. M. (2000). Religion and health: Public health research and practice. *Annual Review of Public Health, 21*, 335–367.

Cohen, S. (2004). Social relationships and health. *The American Psychologist, 59*(8), 676–684.

Cohn, M. A., Fredrickson, B. L., Brown, S. L., Mikels, J. A., & Conway, A. M. (2009). Happiness unpacked: Positive emotions increase life satisfaction by building resilience. *Emotion, 9*(3), 361–368.

Dezutter, J., Soenens, B., & Hutsebaut, D. (2006). Religiosity and mental health: A further exploration of the relative importance of religious behaviors versus religious attitudes. *Personality and Individual Differences, 40*(4), 807–818.

Diener, E., & Chan, M. Y. (2011). Happy people live longer: Subjective well-being contributes to health and longevity. *Applied Psychology: Health and Well-Being, 3*(1), 1–43.

Diener, E., Tay, L., & Myers, D. G. (2011). The religion paradox: If religion makes people happy, why are so many dropping out? *Journal of Personality and Social Psychology, 101*(6), 1278–1290.

Doane, M. J. (2013). The association between religiosity and subjective well-being: The unique contribution of religious service attendance and the mediating role of perceived religious social support. *The Irish Journal of Psychology, 34*(1), 49–66.

Elliott, M., & Hayward, R. D. (2009). Religion and life satisfaction worldwide: The role of government regulation. *Sociology of Religion, 70*(3), 285–310.

Ellison, C. W. (1991). Religious involvement and subjective well-being. *Journal of Health and Social Behavior, 32*, 80–99.

Eriksson, M., & Lindström, B. (2006). Antonovsky's sense of coherence scale and the relation with health: A systematic review. *Journal of Epidemiology and Community Health, 60*, 376–381.

Gadalla, T. M. (2009). Determinants, correlates and mediators of psychological distress: A longitudinal study. *Social Science & Medicine, 68*(12), 2199–2205.

Gale, C. R., Batty, G. D., & Deary, I. J. (2008). Locus of control at age 10 years and health outcomes and behaviors at age 30 years: The 1980 British Cohort Study. *Psychosomatic Medicine, 70*(4), 397–403.

George, L. K., Ellison, C. G., & Larson, D. B. (2002). Explaining the relationships between religious involvement and health. *Psychological Inquiry, 13*(3), 190–200.

Green, M., & Elliott, M. (2010). Religion, health, and psychological well-being. *Journal of Religion and Health, 49*(2), 149–163.

Greenfield, E. A., & Marks, N. F. (2007). Religious social identity as an explanatory factor for associations between more frequent formal religious participation and psychological well-being. *The International Journal for the Psychology of Religion, 17*(3), 245–259.

Grim, B. J., & Finke, R. (2007). Religious persecution in cross-national context: Clashing civilizations or regulated religious economies? *American Sociological Review, 72*(4), 633–658.

Haslam, S., Jetten, J., Postmes, T., & Haslam, C. (2009). Social identity, health and well-being: An emerging agenda for applied psychology. *Applied Psychology: An International Review, 58*(1), 1–23.

Hayward, R., & Elliott, M. (2009). Fitting in with the flock: Social attractiveness as a mechanism for well-being in religious groups. *European Journal of Social Psychology, 39*(4), 592–607.
Koenig, H. G., George, L. K., Blazer, D. G., & Pritchett, J. T. (1993). The relationship between religion and anxiety in a sample of community-dwelling older adults. *Journal of Geriatric Psychiatry, 26*(1), 65–93.
Koenig, H. G., Larson, D. B., & Larson, S. S. (2001). Religion and coping with serious medical illness. *The Annals of Pharmacotherapy, 35*(3), 352–359.
Krause, N., Ellison, C. G., & Wulff, K. M. (1998). Church-based support, negative interaction and psychological well-being: Findings from a national sample of Presbyterians. *Journal for the Scientific Study of Religion, 37*(4), 725–741.
Krause, N., & Wulff, K. M. (2004). Religious doubt and health: Exploring the potential dark side of religion. *Sociology of Religion, 65*(1), 35–56.
Levin, J. (2010). Religion and mental health: Theory and research. *International Journal of Applied Psychoanalytic Studies, 7*(2), 102–115.
Lindenthal, J. J., Myers, J. K., Pepper, M. P., & Stern, M. S. (1970). Mental status and religious behavior. *Journal for the Scientific Study of Religion, 9*(2), 143–149.
Mael, F. A., & Ashforth, B. E. (2001). Identification in work, war, sports, and religion: Contrasting the benefits and risks. *Journal for the Theory of Social Behaviour, 31*(2), 197–222.
Masters, K. S., & Bergin, A. E. (1992). Religious orientation and mental health. In J. F. Schumaker (Ed.), *Religion and mental health* (pp. 221–232). New York: Oxford University Press.
Nicholson, A., Rose, R., & Bobak, M. (2009). Association between attendance at religious services and self-reported health in 22 European countries. *Social Science & Medicine, 69*(4), 519–528.
Pargament, K. I. (1997). *The psychology of religion and coping*. New York: Guilford Press.
Pargament, K. I., Ishler, K., Dubow, E. G., Stanik, P., Rouiller, R., Crowe, P., et al. (1994). Methods of religious coping with the Gulf War: Cross-sectional and longitudinal analyses. *Journal for the Scientific Study of Religion, 33*(4), 347–361.
Pargament, K. I., Smith, B. W., Koenig, H. G., & Perez, L. (1998). Patterns of positive and negative religious coping with major life stressors. *Journal for the Scientific Study of Religion, 37*(4), 710–724.
Pew Research Center. (2012). *"Nones" on the rise: One-in-five adults have no religious affiliation*. Retrieved April 2, 2012, http://www.pewforum.org/uploadedFiles/Topics/Religious_Affiliation/Unaffiliated/NonesOnTheRise-full.pdf.
Schieman, S. (2008). The religious role and the sense of personal control. *Sociology of Religion, 69*(3), 273–296.

# Religion/Spiritual Fulfillment, Satisfaction with

Eduardo Wills-Herrera
School of Management, Universidad de Los Andes, Bogota, DC, Colombia

## Synonyms

Religiousness; Satisfaction with spirituality; Spiritual life; Spiritual socialization; Transcendence; Well-being, spiritual

## Definition

Satisfaction with spiritual fulfillment refers to the ▶ Need of an individual to be part of a greater whole, which in turn influences how the person acts (Erikson, 1982). It has to do with the belief that each individual has a need to contribute to the world in which he or she lives and that such contribution will have a meaning once the person has passed away. Spiritual well-being is a lifelong pursuit and an affirmation of living a life in direct connection with the self, the community, the environment, and the sacred (van Dierendonck & Mohan, 2006). The conceptualization of ▶ Spirituality is particularly challenging due to a wide variety of perspectives and assumptions underlying spirituality research (Sawatzky et al., 2005; Peterson & Seligman, 2004). Spirituality is commonly associated with an existential search for Meaning and Purpose (Larson et al., 1998). Fulfillment with spirituality is an essential field of positive psychology studies (Emmons & Paulatizan, 2003).

Spirituality and religiosity are similar but different concepts. When one speaks of spirituality, one does not need to be affiliated to a specific institution, church, faith, credo, or believe in a particular God, which are more closely linked with organized religion. Some people feel spiritual without necessarily being associated to

a religious institutionalized faith. When one speaks of spirituality, one is referring to matters of the spirit, in particular people's sense of life (Wills, 2009).

## Description

Satisfaction with religion/spiritual fulfillment has been proposed by different trends of research as an important life domain that adds to ▶ satisfaction with life as a whole, ▶ subjective well-being (SWB), ▶ quality of life, and ▶ happiness (Wills, 2009; van Dierendonck, 2011). However, there does not exist a unified definition of what spirituality is and how it can be differentiated from the broader concept of religion that includes a recognizable formal system of beliefs. There is an agreement that spirituality should be defined in the context of individual and cultural experiences (Peterson & Seligman, 2004). It is possible to identify several recent research initiatives that incorporate a spiritual dimension in the measuring of quality of life and well-being despite the fact that it has tended to be neglected in the past because it has been argued that it cannot be scientifically studied (Miller & Thorensen, 2003). In this regard, spirituality has been proposed as a multidimensional variable that includes diverse ▶ values, beliefs, thoughts, experiences, inner resources, and practices. It has also been proposed as a strength and particular character of an individual (Peterson & Seligman, 2004). Among the most cited definitions, spirituality is seen as the experience of a personal relationship with the transcendent or "the belief that each person has a need to contribute to the world in which he lives and that such contribution will have a meaning once the person has passed away" (Erikson, 1982). It has to do with the feeling that an individual is part of a greater whole, a belief which influences how he/she acts. It has been defined also as a "lifelong pursuit and an affirmation of living life in direct connection with self, the community, the environment and the sacred" (van Dierendonck & Mohan, 2006). It has also been treated as a "transcendent dimension that deals with the ultimate goal in life which gives meaning to existence" (Giacalone & Jurkiewickz, 2003). Emmons (1999) has reviewed extensively the role of spiritual satisfaction in relation to well-being, and Paloutzian and Ellison (1982) have proposed the spiritual well-being scale.

Satisfaction with spirituality/religion was proposed as an additional life domain that makes a unique and significant contribution to satisfaction with life as a whole. This fact can be explained by a eudaimonic approach to subjective well-being (Wills, 2009). People evaluate their satisfaction with life as a whole, but they also rate their satisfaction with specific life domains including ▶ health, standard of living, ▶ security, and ▶ community. Research on subjective well-being (Cummins et al., 2003) has attempted to breakdown the global measure of SWB into its constituent components into the ▶ Personal Well-Being Index (PWI). Cummins et al. (2003) have proposed the Personal Well-Being Index as a new scale to assess the minimal set of domains of the construct of personal well-being in order to provide a parsimonious approach of the measure. Once a minimal set of variables is established, a researcher that intends to include a new domain in the PWI should prove empirically that this new domain makes a unique and significant contribution to the prediction of the global measure. Spirituality was tested in a particular cultural context as a new domain because previous studies (Cohen, 2002; Clark & Lelkes, 2005) have proposed that religiosity and spirituality are important considerations for satisfaction with life and that the link between spirituality and life satisfaction should be further researched, particularly among Christians (Protestants and Catholics).

In order to test the association between satisfaction with spirituality and religiosity and satisfaction with life as a whole, a statistical test was performed in a particular cultural context, namely, in Bogota´, the capital city of Colombia. In order to test the contribution of the new domain to PWI, first the empirical validity and ▶ reliability of the PWI and National Well-Being

Index (NWI; Cummins et al., 2003) scales were tested and compared with international results. For this purpose, the study required the translation and back translation of the scale which was originally developed in English for the Australian context (Cummins, 1996). The original English scale of PWI and NWI was translated by the author and compared with similar translations done for Spain.

Respondents were randomly chosen from the total population of the city, which is 7,056,219 inhabitants to have a representative sample by neighborhoods with an error margin of ±4 %. The survey was responded to in an entirely private manner, and respondents were assured that their responses would remain confidential and anonymous.

The seven domains of PWI plus satisfaction with spirituality and religiosity were regressed onto satisfaction with life as a whole (SWLS) (Diener, 1984). All domains with the exception of satisfaction with safety and community connectedness contributed significantly to this regression. Satisfaction with the domains of standard of living, health, achieving in life, and spirituality and religiosity were statistically significant at the $p < 0.00$ and $p < 0.05$ levels. A hierarchical regression analysis with Satisfaction with Life as a whole (Diener et al., 1985) as dependent variable was run. The seven PWI variables were introduced as a first block of independent variables and satisfaction with spirituality and religiosity as a second block of independent variables. It was found that the contribution of the variable spirituality and religiosity in Bogota' was significant, according to its $\Delta R^2$ (0.013) change.

## References


Clark, A. E., & Lelkes, O. (2005). Deliver us from evil: Religion as insurance. PSE Working papers No. 43, PSE, Ecole Normal Superior, Paris.

Cohen, A. B. (2002). The importance of spirituality in well-being for Jews and Christians. *Journal of Happiness Studies, 3*, 287–310.

Cummins, R. A. (1996). The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research, 38*, 303–328.

Cummins, R. A., Eckersley, R., Pallant, J., van Vugt, J., & Misajon, R. (2003). Developing a national index of subjective well-being: The Australian unity well-being index. *Social Indicators Research, 64*, 159–190.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*, 542–575.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71–75.

Emmons, R. A. (1999). *The psychology of ultimate concerns. Motivation and spirituality in personality*. New York: The Guilford Books.

Emmons, R. A., & Paulatizan, R. F. (2003). The psychology of religion. *Annual Review of Psychology, 54*, 377–402.

Erikson, E. (1982). *The life cycle completed: A review*. New York: W.W. Norton.

Giacalone, R. A., & Jurkiewickz, C. L. (2003). *Handbook of workplace spirituality and organizational performance*. New York: Sharpe.

Larson, D. B., Swyers, J. P., & McCullough, M. E. (1998). *Scientific research on spirituality and health: A report based on the scientific progress in spirituality conferences*. New York: The John M. Templeton Foundation.

Miller, W., & Thorensen, C. E. (2003). Spirituality, religion and health: An emerging research field. *American Psychologist, 58*(1), 24–35.

Paloutzian, R. F., & Ellison, C. W. (1982). Loneliness, spiritual well-being, and quality-of-life. In L. A. Peplau & D. Perlman (Eds.), *Loneliness: A sourcebook of current theory, research and therapy* (pp. 224–237). New York: Wiley.

Peterson, C., & Seligman, M. (2004). *Character, strength and virtues. A handbook and classifications*. New York: Oxford University Press.

Sawatzky, R., Ratner, P., & Chiu, L. (2005). A meta-analysis of the relationship between spirituality and quality-of-life. *Social Indicators Research: An International and Interdisciplinary Journal for Quality of Life Measurement, 72*(2), 153–188.

van Dierendonck (2011). Spirituality as an essential determinant for the good life, its importance relative to self-determinant psychological needs. *Journal of Happiness Studies, 13*(4), 685–700.

van Dierendonck, D., & Mohan, K. (2006). Some thoughts on spirituality and eudaimonic well-being. *Mental Health, Religion and Culture, 9*(3), 227–238.

Wills, E. (2009). Spirituality and subjective well-being: Evidences for a new domain in the personal well-being index. *Journal of Happiness Studies, 10*(4), 49–69.


# Religiosity and Chronic Diseases

► Spiritual Needs of Those with Chronic Diseases

## Religiosity and Coping with Schizophrenia

▶ Spirituality and Coping in Patients with Schizophrenia

## Religiosity and Fertility Among US Hispanics

▶ Fertility and Religion Among US Hispanics

## Religiosity and Sexual Attitudes

Tierney Lorenz
Department of Psychology, University of Texas at Austin and University of Washington School of Medicine, Austin, TX, USA

### Definition

#### Religiosity

Religiosity is the degree of importance of religion to an individual. This construct is intuitive but very tricky to measure. Most studies of religiosity include some measure of frequency of participation in religious services (e.g., church attendance). Some also include assessment of the individuals' subjective experience of religion, including the degree to which religion is the organizing principle of everyday life (intrinsic religiosity; Allport, 1950), the strength of connection to a higher power or truth (▶ spirituality), or degree to which one follows religious doctrine dogmatically (fundamentalism). Much of the research on religiosity and sexual ▶ attitudes has been conducted in predominantly Christian countries. Thus, although this entry includes findings from non-Christian religions when possible, if not otherwise indicated, it should be assumed that "religiosity" refers to Christian religiosity.

#### Sexual Attitudes

A sexual attitude is a tendency in thoughts and emotions related to sexuality. In academic literature, the term "positive sexual attitudes" generally signifies more liberal or permissive attitudes towards free sexual expression, sex as recreational not (necessarily) procreative, and so on.

### Description

#### Sexual Attitudes as Predictors of Sexual Behavior

Sexual attitudes are thought to be the means by which religion regulates sexual behavior; however, several studies have suggested other intermediaries such as sexual guilt (Woo, Morshedian, Brotto, & Gorzalka, 2012). As with all attitudes, there is a difference between sexual attitudes and sexual behavior. Highly religious individuals tend to report more conservative sexual behavior, including later age of first sex (Thornton & Camburn, 1989), fewer sexual partners (Davidson, Moore, & Ullstrup, 2004), and less variation of sexual practices (e.g., less oral sex; Regenerus, 2007). This dampening effect of religiosity on sexual behavior is mediated through sexual attitudes (Meier, 2003). However, religiosity does not overwhelmingly override other influences on sexual behavior such as peer group attitudes (Burdette & Hill, 2009). Interestingly, individuals reporting high fundamental religiosity alongside low spirituality tend to have greater sexual attitude-behavior inconsistency (Ahrold & Meston, 2007). This may be because while sexual behavior changes sexual attitudes (e.g., first sex tends to liberalize sexual attitudes), it does not change religiosity (Meier, 2003).

#### Aspects of Religiosity and Sexual Attitudes

Many studies have investigated the impact of individual elements of religiosity on sexual attitudes, often with conflicting results. However, across studies, intrinsic religiosity appears to predict attitudes towards social regulation of sexuality (e.g., attitudes towards premarital sex). Fundamentalism similarly predicts sexual attitudes related to interpersonal sexuality (Hunsberger, 2010),

but additionally predicts attitudes towards aspects of sexuality related to the self (e.g., masturbation, attitudes towards sexual fantasy; Ahrold, Farmer, Trapnell, & Meston, 2011; Davidson et al., 2004).

This difference may be due to the factors underlying intrinsic religiosity and fundamentalism. Intrinsic religiosity is very closely associated with church attendance (Zinnbauer et al., 1997). Insofar as intrinsic religiosity is gained through socialization in religious organizations, it orients the individual towards the social function of religion (Zinnbauer et al., 1997). Simply put, if you think of yourself as very religious because you spend a lot of time engaged in religious practices with others (i.e., high religiosity), you likely think more about how your religion directs you to act among others than if you thought of yourself as very religious because you have a close relationship with God (e.g., high spirituality). Church attendance is a necessary but not sufficient predictor of fundamentalism: dogmatism is bred not only by socialization but individual orientation towards conservatism. In other words, fundamentalism may arise from a rigid view of religion applied equally to self and society.

The dimensions of religiosity may interact to produce different effects than each separately. For example, women who report high intrinsic religiosity tend to have more conservative attitudes towards extramarital sexuality, ▶ casual sex, and homosexuality; however, this effect is greatly dulled if women report low spirituality (Ahrold & Meston, 2010). Similarly, although they differ in religious participation, intrinsic religiosity, and fundamentalism each separately predict conservative sexual attitudes, when considered together (i.e., in a regression in which all covariance is partialed out), only religious participation and to a much lesser extent intrinsic religiosity remain significant predictors (Lefkowitz, Gillen, Shearer, & Boone, 2004).

## Secularization Versus Religious Transformation

In the Western world, religion has shifted from cultural hegemony to one of many influences on attitudes. As individuals in a society gain access to diverse viewpoints and information, religious institutions have a less unique influence in defining those individuals' identities and, thus, their sexual attitudes (Stark & Bainbridge, 1985). This shift in the role of religion has been termed "secularization" and has been the presumed mechanism through which sexual attitudes change at a societal level: as society becomes more secular, it is thought to become more sexually liberal (Farmer, Trapnell, & Meston, 2009).

Those religions that have thrived in the Western world do so by "transforming": accommodating individual choice to engage in religion as well as coexistence with other influences (e.g., of scientific thought). This adaptive process has been termed "religious transformation." Ahrold et al. (2011) showed that when controlling for spirituality – the perceived strength of individual relationships with the Divine – religious group differences in conservative sexual attitudes drop out. This supports the religious transformation hypothesis: while religion may play a lesser role in regulation of sexuality in society as a whole, it continues to influence sexual attitudes among the religious (Ahrold et al., 2011).

## Attitudes Towards Nonmarital Sexuality

Because a great deal of scientific research on sexual attitudes has been conducted in adolescents and young adults (see Regenerus, 2007, for review; two notable exceptions are the National Health and Social Life Survey in the USA and the National Survey of Sexual Attitudes and Lifestyles in the UK), research on nonmarital sexuality has been a focus on attitudes towards premarital sexuality (and not adultery or consensually nonmonogamous relationships). Not surprisingly, considering a central aspect of most religions revolves around upholding marriage (however that religion defines it), religiosity tends to predict positive attitudes towards marriage and negative attitudes towards sex outside the confines of marriage (although religions differ in the degree to which procreation is the intended outcome of marriage; see below).

Higher religiosity predicts more conservative attitudes towards premarital sexuality (Lefkowitz et al., 2004), but when respondents are specifically asked about couples planning to be married, these

effects drop out (Thornton & Camburn, 1989). Among adolescents and young adults, women report higher intrinsic religiosity than do men (Milevsky & Levitt, 2004) and also report less acceptance of sexuality outside the context of long-term committed relationships (Oliver & Hyde, 1993). Most research on this gender effect has attempted to demonstrate that women tend to be more religious and thus less accepting of nonmarital sexuality (e.g., Lefkowitz et al., 2004); however, it is also possible that women tend to have more conservative beliefs about sexuality (and other societal norms) which make them more comfortable in a religious environment.

### Attitudes Towards Homosexuality

With a few exceptions, higher religiosity tends to be associated with homophobia (see linked article for an in-depth discussion of the impact of religion on attitudes towards homosexuality). In Judeo-Christian communities, negative attitudes towards male-male sexuality are generally stronger than attitudes towards female-female sexuality (Herek, 1987), while in Hindu and Muslim communities, this divergence is less pronounced (Adamczyk & Pitt, 2009).

### Attitudes Towards Contraception

Although church attendance has been shown to be a protective factor against sexually transmitted infections by way of later age of sexual debut and fewer sexual partners (Barkan, 2006), it is also associated with more negative attitudes towards contraception (Lefkowitz et al., 2004). One notable exception is that in African American communities, higher intrinsic religiosity predicts more positive attitudes towards condom use (McCree, Wingood, DiClemente, Davies, & Harrington, 2003). Gender differences have also been reported with higher religiosity associated with more positive attitudes towards contraceptives in adolescent females, but more negative attitudes in males (Manlove, Logan, Moore, & Ikrawmullah, 2008). Of course, religiosity is a stronger predictor of negative attitudes towards contraceptives for members of religions that officially denounce use of contraceptives (e.g., Catholicism) than for religions that are neutral or encourage contraceptive use (e.g., Reform Judaism; Weisberg & Kern, 2009).

### Prosexual Attitudes and Religiosity

It is often assumed that religion is at cross-purposes with prosexual attitudes (see Regenerus (2007) for in-depth discussion). Certainly, several studies have shown that higher religiosity is associated with higher sexual guilt in Christians (Wyatt & Dunn, 1991) and Buddhists (Woo et al., 2012). However, religious participation is not significantly related to positive attitudes about one's own sexual relationship or ▶ sexual satisfaction (Davidson et al., 2004). Similarly, religiosity is positively associated with higher positive attitudes about marriage and ▶ relationship satisfaction (Mahoney et al., 1999). Certain types of religiosity, such as belief in paranormal or supernatural powers, are associated with more liberal attitudes towards masturbation (Ahrold et al., 2011).

Judaism can be seen as having decidedly prosexual attitudes, as sex (between married people and at appropriate times of a woman's menstrual cycle) is seen as a mitzvah, or a good deed done in the service of religious duty, even in the absence of procreation. Jews tend to have more liberal sexual attitudes, particularly towards contraception and abortion, than other religions (Leiblum, Weigel, & Brickle, 2003; Notzer, Levran, Mashiach, & Sqffer, 1984; Schenker, 2000; Weis & Slosnerick, 1981); this difference is small but consistent across studies.

Both Buddhism and Hinduism have ambivalent sexual attitudes: both negative, with doctrine encouraging sexual restraint by way of avoiding over attachment to worldly desires (Sobo & Bell, 2001), and positive, with ancient Taoist and Tantric practices supported sexuality as a means of accessing the Divine (Chang, 2001). The impact of Buddhism on conservative sexual attitudes is difficult to distinguish from the effect of Asian ethnicity (Ahrold & Meston, 2010).

Within the Christian tradition, the United Church of Christ has notably prosexual attitudes, encouraging practitioners to engage in lifelong comprehensive sex education (Kimball, 2000) and formally supporting same-sex marriage

("Equal marriage rights for all," 2005) and abortion rights ("Sexuality and abortion: A faithful response," 1987); these are echoed by the Unitarian Universalists, who are derived from a Judeo-Christian tradition. Interestingly, several prominent Christian apologists have argued that the sexual attitudes presented in the New Testament are generally neutral and far less negative than the attitudes towards intimacy between mortals. C. S. Lewis writes in his essay *The Four Loves* (1960),

> St. Paul, dissuading his converts from marriage, says nothing about that side of the matter except to discourage prolonged abstinence from *(the sexual act)*. What he fears is pre-occupation, the need of constantly pleasing... one's partner, the multiple distractions of domesticity. It is marriage itself, not the marriage bed, that will be likely to hinder us from waiting uninterruptedly on God.

## References


Adamczyk, A., & Pitt, C. (2009). Shaping attitudes about homosexuality: The role of religion and cultural context. *Social Science Research, 38*(2), 338–351.

Ahrold, T. K., Farmer, M., Trapnell, P., & Meston, C. M. (2011). The relationship among sexual attitudes, sexual fantasy, and religion. *Archives of Sexual Behavior, 40*, 619–630.

Ahrold, T. K., & Meston, C. M. (2007, August). Sure it's fine, but I don't do it: Gender differences in attitude/behavior inconsistency in masturbation. *Poster presented at the meeting of the International Association for Sex Research (IASR)*. Vancouver, Canada.

Ahrold, T. K., & Meston, C. M. (2010). Ethnic differences in sexual attitudes of U.S. college students: Gender, acculturation, and religiosity factors. *Archives of Sexual Behavior, 39*, 190–202.

Allport, G. (1950). *The individual and his religion*. New York: Macmillan.

Barkan, S. E. (2006). Religiosity and premarital sex in adulthood. *Journal for the Scientific Study of Religion, 45*(3), 407–417.

Burdette, A., & Hill, T. D. (2009). Religious involvement and transitions into adolescent sexual activities. *Sociology of Religion, 70*(1), 5–27.

Chang, S. (2001). *The Tao of sexology*. London: Tao.

Davidson, K., Moore, N. B., & Ullstrup, K. M. (2004). Religiosity and sexual responsibility: Relationships of choice. *American Journal of Health Behavior, 28*(4), 335–346.

Equal marriage rights for all. (2005, July). Resolution passed by the 25th Synod of the United Church of Christ, Cleveland, OH.

Farmer, M. A., Trapnell, P. D., & Meston, C. M. (2009). The relation between sexual behavior and religiosity subtypes: A test of the secularization hypothesis. *Archives of sexual behavior, 38*(5), 852–865.

Herek, G. (1987). Religious orientation and prejudice: A comparison of racial and sexual attitudes. *Personality and Social Psychology Bulletin, 13*(1), 33–44.

Hunsberger, B. (2010). Religion and Prejudice: The role of religious fundamentalism, quest, and right-wing authoritarianism. *Journal of Social Issues, 51*(2), 113–129.

Kimball, R. (2000). *Our whole lives: Sexuality education for adults*. Boston: Unitarian Universalist Association.

Lefkowitz, E., Gillen, M., Shearer, C., & Boone, T. (2004). Religiosity, sexual behaviors, and sexual attitudes during emerging adulthood. *Journal of Sex Research, 41*(2), 150–159.

Leiblum, S., Weigel, M., & Brickle, F. (2003). Sexual attitudes of US and Canadian medical students: The role of ethnicity, gender, religion and acculturation. *Sexual and Relationship Therapy, 18*(4), 473–491.

Lewis, C. S. (1960). *The four loves*. New York: HarperCollins.

Mahoney, A., Pargament, K. I., Jewell, T., Swank, A. B., Scott, E., Emery, E., et al. (1999). Marriage and the spiritual realm: The role of proximal and distal religious constructs in marital functioning. *Journal of Family Psychology, 13*, 321–338.

Manlove, J., Logan, C., Moore, K., & Ikrawmullah, E. (2008). Pathways from family religiosity to adolescent sexual activity and contraceptive use. *Perspectives on Sexual and Reproductive Health, 40*(2), 105–117.

McCree, D. H., Wingood, G. M., DiClemente, R., Davies, S., & Harrington, K. F. (2003). Religiosity and risky sexual behavior in African-American adolescent females. *Journal of Adolescent Health, 33*(1), 2–8.

Meier, A. (2003). Adolescents' transition to first intercourse, religiosity, and attitudes about sex. *Social Forces, 81*(3), 1031–1052.

Milevsky, A., & Levitt, M. (2004). Intrinsic and extrinsic religiosity in preadolescence and adolescence: Effect on psychological adjustment. *Mental Health, Religion and Culture, 7*(4), 307–321.

Notzer, N., Levran, D., Mashiach, S., & Sqffer, S. (1984). Effect of religiosity on sex attitudes, experience and contraception among university students. *Journal of Sex & Marital Therapy, 10*(1), 57–62.

Oliver, M., & Hyde, J. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin, 114*(1), 29–51.

Regenerus, M. (2007). *Forbidden Fruit? Sex and religion in American adolescence*. New York: Oxford University Press.

Schenker, J. G. (2000). Women's reproductive health: Monotheistic religious perspectives. *International Journal of Gynecology & Obstetrics, 70*(1), 77–86.

Sexuality and abortion: A faithful response. (1987, July). Resolution of the 16th Synod of the United Church of Christ, Cleveland, OH.

Sobo, E., & Bell, S. (2001). *Celibacy, culture, and society: Anthropology of sexual abstinence*. Madison, WI: University of Wisconsin Press.
Stark, R., & Bainbridge, W. S. (1985). *The future of religion: Secularization, revival and cult formation*. Berkeley, CA: University of California Press.
Thornton, A., & Camburn, D. (1989). Religious participation and adolescent sexual behavior and attitudes. *Journal of Marriage and the Family*, 641–653.
Weis, D. L., & Slosnerick, M. (1981). Attitudes toward sexual and nonsexual extramarital involvements among a sample of college students. *Journal of Marriage and the Family*, 349–358.
Weisberg, E., & Kern, I. (2009). Judaism and women's health. *The Journal of Family Planning and Reproductive Health Care, 35*(1), 53–55.
Woo, J. S. T., Morshedian, N., Brotto, L. A., & Gorzalka, B. B. (2012). Sex guilt mediates the relationship between religiosity and sexual desire in East Asian and Euro-Canadian college-aged women. *Archives of Sexual Behavior, 41*, 1485–1495.
Wyatt, G., & Dunn, K. (1991). Examining predictors of sex guilt in multiethnic samples of women. *Archives of Sexual Behavior, 20*(5), 471–485.
Zinnbauer, B. J., Pargament, K. I., Cole, B., Rye, M. S., Butter, E. M., Belavich, T. G., et al. (1997). Religion and spirituality: Unfuzzying the fuzzy. *Journal for the Scientific Study of Religion, 36*(4), 549–564.

# Religiosity and Support for Democracy

Pazit Ben-Nun Bloom
Department of Political Science, The Hebrew University of Jerusalem, Jerusalem, Israel

## Definition

The role of religion in democratic ▶ attitudes has been the focus of countless theories and studies, with some scholars arguing that religion has the effect of challenging democratic values and may thus pose an obstacle to political socialization that shapes these values and other scholars stressing its contribution to constituents' civic skills, institutional ▶ trust, and support for a democratic regime. How can these contradictory views be reconciled? It was recently suggested that a way of settling this debate lies in viewing religiosity as a multifaceted phenomenon, with different dimensions exerting different, and even contrasting, effects on democratic attitudes.

## Description

▶ Democracy is not simply attained by making the necessary institutional changes; its survival also depends on the ▶ values and beliefs of ordinary citizens (Inglehart, 2000). Indeed, the consolidation of democracy seems to be more successful where mass commitment to democratic ▶ norms and institutions is high (Chu, Diamond, & Shin, 2000). Social scientists have thus devoted much research to identifying the individual-level correlates that may affect the development of democratic attitudes, with religiosity being a prime suspect.

Starting with Adorno, Frenkel-Brunswik, Levinson, and Sanford's (1950) theory of the authoritarian personality, which suggested a psychoanalytical connection between tendencies toward religious and undemocratic attitudes, many scholars have argued that religion has the effect of challenging democratic values and socialization (Stark, 2001). Much of this literature has contrasted religion and democracy as systems of belief and has focused on the challenge religious extremism and loyalties pose for democratic institutions. Religion is often presented as undermining some human freedoms and civil liberties such as ▶ gender equality and minority rights. Additionally, democratic values are argued to stress universality, striving for global implementation of civil rights for every person, whereas the religious public typically considers itself as superior to other groups and usually entitled to more rights than others (e.g., Ben-Nun Bloom, Arikan, & Sommer, 2013). Indeed, religiosity was found connected to such proclivities as political intolerance, ▶ prejudice, extremity, and nondemocratic norms (Canetti-Nisim, 2004; Gibson, 1992; Hunsberger, 1995; Karpov, 2002; Stark, 2001).

However, at the same time, religiosity was found to have a positive impact on democratic norms and values. Evidence shows that churches

hold great potential for deliberative democracy (Neiheisel, Djupe, & Sokhey, 2008), as religious activity leads to the development of civic skills and civic norms, and provides organizational and philosophical bases for a wide range of social movements (Putnam, 2000). In fact, church attendance is found to increase electoral turnout, party membership, protest activism, and support for democracy (Norris & Inglehart, 2004; Meyer, Tope, & Price, 2008). Accordingly, empirical and theoretical scholars alike increasingly note the "political ambivalence of religion" (Appleby, 2000; Bloom, 2012 ; Philpott, 2007), that is, the fact that it can be either a source of undemocratic values or a contributor to the development of democratic skills.

These contradictory findings can be reconciled when considering the multifaceted nature of religiosity. The literature views religion as involving three dimensions – *belief*, *behavior*, and *belonging* (Guth, Jelen, Kellstedt, Smidt, & Wald, 1988; Kellstedt, Green, Guth, & Smidt, 1997; Smidt, Kellstedt, & Guth, 2009). The *belief* component encompasses theology, "an understanding of the divine and humanity's relationship to it," and "social theology," which "connects the individual and the institutional church to the world" (Kellstedt et al., 1997), and may refer to belief in God, heaven, hell, life after death, or tendencies of people to characterize themselves as religious (Barro & McCleary, 2003). The *behavior* component consists of two factors: the social practice of religion, involving participation in organized religious communities and attendance at places of worship, and private practice such as prayer. *Belonging* consists of denominational affiliation, that is, identification with a particular organized denomination and/or a religious movement, and private practice, such as prayer or the reading of sacred texts. The belonging component includes identification with a specific denomination and trends within a denomination (Kellstedt et al., 1997).

Accordingly, current literature argues that religious beliefs, such as the belief in God, heaven, and life after death, and social religious activities, such as attendance at places of worship and participation in organized religious communities, often have contrasting effects on democratic attitudes and norms (Ben-Nun Bloom & Arikan, 2012, 2013a, b). First, religious belief is positively associated with conservative-traditional values and negatively associated with openness to change values across religions and contexts (see Saroglou, Delpierre, & Dernelle, 2004 for a meta-analysis; and Lau, 1989; Rokeach, 1968; Schwartz & Huismans, 1995), whereas democratization is positively related to openness to change, emphasizing independent thought, universalism, natural rights, and ▶ equality, but negatively related to values such as conformity, tradition, and security (Schwartz & Sagie, 2000). This makes an inherent and systematic value conflict between the religious and democratic values systems. At the same time, social involvement in places of worship leads to the development of civic skills and norms as well as ▶ political efficacy (Putnam, 2000); positively affects electoral turnout, party membership, protest activism, and involvement in other civic organizations (Norris & Inglehart, 2004); and thus holds great potential for deliberative democracy (Neiheisel et al., 2008). Further, the religious social institution makes for an active minority group that benefits from the democratic framework, consequently mobilizing overall support for a democratic regime.

Indeed, using heteroskedastic maximum likelihood models and data from the fourth wave of the World Values Survey for 45 democratic countries, Ben-Nun Bloom and Arikan (2012) show that as a belief system, religiosity generates abstract opposition to democracy while increasing ambivalence toward democratic principles. However, the social behavioral aspect of religiosity leads to greater endorsement of the democratic system and generates stronger support for democracy. Theoretically, this work argues that different mediating mechanisms underlie the differential effects that the two dimensions of religiosity have on attitudes toward democracy; while religious belief is associated with traditional values and an emphasis on material and physical security that generates opposition to democracy, religious social behavior improves social capital in the form of institutional trust

and ▶ political engagement, which have positive effects on support for democracy. In a study that used multilevel path analysis models and data from 54 countries from Waves 4 and 5 of the World Values Survey to directly test the mechanisms hypothesized to underlie the differential effects of religious belief and behavior on abstract support for democracy, it was found that the negative effect of religious belief on democratic attitudes is to a large extent mediated by personal values and the effect of social religious behavior is mediated by the generation of social capital in the form of political involvement and trust in institutions (Ben-Nun Bloom & Arikan, 2013a).

While these studies confirm that religiousness affects democratic support through values, involvement, and trust in democratic institutions, it could be argued that democratic norms and attitudes increase the likelihood of political participation via places of worship and not the other way around. Similarly, there could be some authoritarian or otherwise conservative trait that is simultaneously responsible for both religious belief and undemocratic attitudes. Therefore, to improve our understanding of the nature of the relationship between the two dimensions of religiosity and democratic attitudes, it is necessary to experimentally establish their causal effect.

Until recently, ▶ experiments manipulating religion were relatively rare, as levels of religiosity do not appear to be variables that can be varied in the lab. However, recent multidisciplinary literature builds on a priming framework to experimentally study the effects of religion. A priming framework suggests that the subject's exposure to certain cues increases the accessibility of related objects in his memory, consciously or unconsciously (Krosnick & Kinder, 1990). Thus, even relatively subtle religious cues can activate religious beliefs or values and experiences, which, in turn, can affect political cognition. Indeed, primes and cues on religion were found to have an effect on pro-social attitudes, moral behavior, prejudice, and candidate evaluation (McKay, Efferson, Whitehouse, & Fehr, 2011; Pichon, Boccato, & Saroglou, 2007; Shariff & Norenzayan, 2007). Still, current experimental studies typically confound the two dimensions of religiosity (but see Ginges, Hansen, & Norenzayan, 2009).

Therefore, to test the causal effect of religious belief and religious social behavior on abstract support for democracy, a recent study builds on an experimental priming framework. Using a comparative experiment among Turkish Muslims and Israeli Jews, it was found that priming religious belief led to a lower degree of abstract support for democracy among respondents, while a religious social behavior prime led to an increase in abstract support for democracy when compared to the no-prime group (Ben-Nun Bloom & Arikan, 2013b).

While current results robustly confirm the differential effects of faith and religious social identity on support for democracy, some future studies are still in order. Public demand for democracy is specifically important in authoritarian regimes and in systems moving toward a ▶ liberal democracy, as it may affect whether an open democratic society will develop or a relapse into a more authoritarian mode will occur (Welzel, 2007). At the same time, large segments of the population in many nondemocratic regimes and nascent democracies are religiously orthodox, making the role of religiosity in affecting public demand for liberal democracy particularly important in these cases. Yet, current empirical studies simultaneously testing the effect of different dimensions of religiosity focus on support for democracy in established or emerging democracies, leaving room for future studies of the multidimensional effects of religion on democratic demand in nondemocracies.

Next, the above-described works regarding the differential effects of different dimensions of religiosity mostly focus on testing the effect of religiosity on "abstract" or "overt" support for democracy, that is, citizens' declaration of general endorsement of the democratic regime and of deeming it desirable for their country. While such declarative support is often viewed as critical for the thriving of democratic culture (Almond and Verba, 1963; Dalton, 2004; Easton, 1975; Inglehart 2003), democracy is a multidimensional concept including substantive support for specific regime components, such as endorsement of democratic

procedures, values, and norms, including free elections, civil rights, gender equality, and freedom of speech (Dalton, 1999; Easton, 1975; Norris, 2011), as well as rejection of autocratic principles and instrumental outcomes that are *not* integral to liberal democracy, such as economic prosperity, relative security, lower crime levels, and redistribution of ▶ wealth (Klingemann, 1999; Norris, 1999; 2011). Consequently, future research should integrate the current literature on the multidimensionality of both religiosity and democratic support.

Finally, the religious ministry may shape followers' approach to democracy by communicating messages regarding the desirable relationship between religious authorities and the state and the obligations of the devout toward the political authority (Philpott, 2007). Thus, the effect of social religious behavior on democratic support may be dependent on the informational environment within which the individual is located. Religious elites may act to intensify the conflict between religion and democracy – either because they perceive a vast conflict between the two belief systems or for some political reason – and as a consequence their followers are more likely to decrease in democratic support; alternatively, religious elites may try to eliminate the conflict, by stressing that the two belief systems go hand in hand, and thereby contribute to democratic support among their adherents (Alvarez and Brehm, 1995:1077). Whether the effect of religiosity on democratic attitudes is positive or negative, then, may be affected by elite messages regarding the conflict between religion and democracy. Therefore, future studies should investigate the role of religious elites in channeling the value conflict experienced by their constituency.

## References


Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, N. R. (1950). *The authoritarian personality*. New York: Harper and Row.

Almond, G. A., & Verba, S. (1963). *The civic culture: Political attitudes and democracy in five nations*. Princeton, NJ: Princeton University Press.

Alvarez, R. M., & Brehm, J. (1995). American ambivalence toward abortion policy: Development of a heteroskedastic probit model of competing values. *American Journal of Political Science, 39*, 1055–1082.

Appleby, S. R. (2000). *The ambivalence of the sacred: Religion, violence, and reconciliation*. New York: Rowman and Littlefield.

Barro, R. J., & McCleary, R. M. (2003). Religion and economic growth across countries. *American Sociological Review, 68*(5), 760–781.

Ben-Nun Bloom, P., & Arikan, G. (2012). A two-edged sword: The differential effect of religious belief and religious social context on attitudes towards democracy. *Political Behavior, 34*(2), 249–276.

Ben-Nun Bloom, P., & Arikan, G. (2013a). Religious belief, religious social behavior, and democratic attitudes: A cross-country test of the mediating mechanisms. *British Journal of Political Science* (in press). http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8698011

Ben-Nun Bloom, P., & Arikan, G. (2013b). Priming religious belief and religious social behavior affects support for democracy. *International Journal of Public Opinion Research* (in press). http://ijpor.oxfordjournals.org/content/early/2012/11/18/ijpor.eds030.short?rss=1

Ben-Nun Bloom, P., Arikan, G., & Sommer, U. (2013). Globalization, threat and religious freedom. *Political Studies* (in press).

Bloom, P. (2012). Religion, morality, evolution. *Annual Review of Psychology, 63*, 179–199.

Canetti-Nisim, D. (2004). The effect of religiosity on endorsement of democratic values: The mediating influence of authoritarianism. *Political Behavior, 26*(4), 377–398.

Chu, Y., Diamond, L. & Shin, D.C. (2000). Growth and equivocation in support for democracy: Taiwan and Korea in comparative perspective. *Paper presented at the annual meeting of the American Political Science Association*, Washington, D.C., August 31-September 3, 2000.

Dalton, R. J. (1999). Political support in advanced industrial democracies. In P. Norris (Ed.), *Critical citizens: Global support for democratic governance*. Oxford: Oxford University Press.

Dalton, R. J. (2004). *Democratic challenges, democratic choice: The erosion of political support in advanced industrial democracy*. New York: Oxford University Press.

Easton, D. (1975). A re-assessment of the concept of political support. *British Journal of Political Science, 5*, 435–457.

Gibson, J. L. (1992). The political consequences of intolerance: Cultural conformity and political freedom. *American Political Science Review, 86*(2), 338–356.

Ginges, J., Hansen, I., & Norenzayan, A. (2009). Religion and support for suicide attacks. *Psychological Science, 20*(2), 224–230.

Guth, J. L., Jelen, T. G., Kellstedt, L. A., Smidt, C. E., & Wald, K. D. (1988). The politics of religion in America: Issues for investigation. *American Politics Quarterly, 16*(3), 357–397.
Hunsberger, B. (1995). Religion and prejudice: The role of religious fundamentalism, quest and right-wing authoritarianism. *Journal of Social Issues, 51*(2), 113–129.
Inglehart, R. F. (2000). Culture and democracy. In L. E. Harrison & S. Huntington (Eds.), *Culture matters: How values shape human progress*. New York: Basic Books.
Inglehart, R. F. (2003). How strong is mass support for democracy -and how can we measure it? *Political Science and Politics, 36*(1), 51–57.
Karpov, V. (2002). Tolerance in the United States and Poland. *Journal for the Scientific Study of Religion, 41*(2), 267–288.
Kellstedt, L.A., Green, J.C., Guth, J.L. & Smidt, C.E. (1997). Is there a culture war? Religion and the 1996 election. *Paper presented at the Annual Meeting of the American Political Science Association*, August 28–31, Washington, DC.
Klingemann, H. D. (1999). Mapping political support in the 1990s: A global analysis. In P. Norris (Ed.), *Critical citizens: Global support for Democratic governance* (pp. 31–56). New York: Oxford University Press.
Krosnick, J. A., & Kinder, D. R. (1990). Altering the foundations of support for the President through priming. *American Political Science Review, 84*(2), 497–512.
Lau, S. (1989). Religious schema and values. *International Journal of Psychology, 24*, 137–156.
McKay, R., Efferson, C., Whitehouse, H., & Fehr, E. (2011). Wrath of God: Religious primes and punishment. *Proceedings of the Royal Society B: Biological Sciences, 278*(1713), 1858–1863.
Meyer, K., Tope, D., & Price, A. M. (2008). Religion and support for democracy: A cross-national examination. *Sociological Spectrum, 28*(5), 625–653.
Neiheisel, J. R., Djupe, P. A., & Sokhey, A. E. (2008). Veni, vidi, disseri: Churches and the promise of democratic deliberation. *American Politics Research, 20*(10), 1–31.
Norris, P. (1999). Institutional explanations for political trust. In P. Norris (Ed.), *Critical citizens: Global support for democratic governance* (pp. 217–235). Oxford, UK: Oxford University Press.
Norris, P., & Inglehart, R. (2004). *Sacred and secular: Religion and politics worldwide*. Cambridge, UK/New York: Cambridge University Press.
Norris, P. (2011). *Democratic deficit: Critical citizens revisited*. Cambridge, UK/New York: Cambridge University Press.
Philpott, D. (2007). Explaining the political ambivalence of religion. *American Political Science Review, 101*(3), 505–525.
Pichon, I., Boccato, G., & Saroglou, V. (2007). Nonconscious influence of religion on prosociality: A priming study. *European Journal of Social Psychology, 37*(January), 1–6.
Putnam, R. D. (2000). *Bowling alone*. New York: Simon and Schuster.
Rokeach, M. (1968). *Beliefs, attitudes and values*. San Francisco, CA: Jossey-Bass.
Saroglou, V., Delpierre, V., & Dernelle, R. (2004). Values and religiosity: A meta-analysis of studies using Schwartz's model. *Personality and Individual Differences, 37*(4), 721–734.
Schwartz, S., & Huismans, S. (1995). Value priorities and religiosity in four Western religions. *Social Psychology Quarterly, 58*(1), 88–107.
Schwartz, S. H., & Sagie, G. (2000). Value consensus and importance: A cross-national study. *Journal of Cross-Cultural Psychology, 31*(4), 465–497.
Shariff, A. F., & Norenzayan, A. (2007). God is watching you. *Psychological Science, 18*(9), 803–809.
Smidt, C. E., Kellstedt, L. A., & Guth, J. L. (Eds.). (2009). The role of religion in American politics: Explanatory theories and associated analytical and measurement issues. *Oxford handbook on religion and American politics (pp. 3–42)*. Oxford: Oxford University Press.
Stark, R. (2001). Reconceptualizing religion, magic, and science. *Review of Religious Research, 43*(2), 101–120.
Welzel, C. (2007). Are levels of democracy affected by mass attitudes? Testing attainment and sustainment effects on democracy. *International Political Science Review, 28*(4), 397–424.

## Religiosity and Well-Being

▶ Religion, Psychological Well-Being, and Health



## Religiosity for Muslim Students in Kuwait, Quality of Life

Ahmed M. Abdel-Khalek
Department of Psychology, Faculty of Arts, University of Alexandria, Alexandria, Egypt

### Definition

There is a classical distinction between objective and subjective QOL. The objective aspect includes characteristics of the environment,

► education, ► health system, safety, opportunities, income, and facilities, among others. The subjective QOL incorporates the satisfaction with the last-mentioned items and the self-appraisal of the living conditions, as well as feelings of ► happiness. However, Veenhoven (2011) criticized this approach and proposed a fourfold matrix classification of the QOLs, i.e., qualities not a quality. His schema includes four quadrants indicating four QOLs as follows: (1) ► livability of the environment, (2) life ability of the person, (3) external utility of life, and (4) subjective ► enjoyment of life. He concluded that the most comprehensive and inclusive summary measure for QOL is how long and happily a person or people live(s).

The World Health Organization (WHO) defined QOL as:

> Individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the persons' physical health, psychological state, level of independence, social relationships and their relationship to salient features of their environment. (WHOQOL Group, 1995, p. 1404).

## Description

### Religiosity

► Religion is considered to be one of the major influential forces across history. There is reason to believe that religion has returned back to its predecessor's high status. Novak (1998) stated that the twenty-first century will be "the most religious century" in recent years.

Thoresen (1998) defined religion as an organized system of beliefs, practices, rituals, and symbols designed to facilitate a relationship to, and an understanding of, a deity as well as to promote the understanding and harmony of a person's relationship to oneself and others. The same author defined religiosity as a person's adherence to the beliefs, values, and practices proposed by an organized institution, which is devoted to the search for the divine through prescribed ways of viewing and living life.

Koenig, McCullough, and Larson (2001) described religiosity as a system of beliefs and practices that may include both internal and external forms of religious activity, such as prayer or attendance at religious service. Such practices are designed to foster closeness to the sacred and an understanding of one's relationships. Allport (1959) elaborated the difference between extrinsic religiosity (the outward signs of religious socialization, such as churchgoing) and intrinsic religiosity (inward depth of feeling). He concluded that they were two distinct parallel phenomena. This differentiation has won widespread acceptance. Thorson, Powell, Abdel-Khalek, and Beshai (1997) found that Kuwaiti undergraduates, exclusively Muslims, were significantly more religious than did their American counterparts.

### The Relationship Between QOL and Religiosity

Ferriss (2002) examined the relationship between religion and the quality of life in its two indicators, i.e., the objective measures including longevity as an indicator of the ► good life, and the proportion of the population professing a religious faith, and the subjective measure, namely, happiness. He stated an important study by demographers that has established a firm relationship between length of life and religion. He also documented positive relations between aspects of religiosity and happiness and the satisfaction with life.

A meta-analytical study of the relationship between spirituality and quality of life was carried out by Sawatzky, Ratner, and Chiu (2005). They included in the final analysis 62 primary effect sizes from 51 studies. They found a moderate effect size ($r = 0.34$, 95 % CI: 0.28–0.40) and concluded that the results of the meta-analysis support the conceptualization of spirituality as a distinct concept that relates to quality of life. Zullig, Ward, and Horn (2006) stated that several studies have suggested that religiosity and spirituality have a significant effect on measures of QOL, even when controlling for sociodemographic variables.

A large individual data set from the Eurobarometer survey used by Greene and

**Religiosity for Muslim Students in Kuwait, Quality of Life, Table 1** Mean (*M*), standard deviation (*SD*), *t* value, and the significance level (*p*) among Kuwaiti men ($n = 106$) and women ($n = 118$)

| | Men | | Women | | | |
|---|---|---|---|---|---|---|
| Scales | *M* | *SD* | *M* | *SD* | *t* | *p* |
| WHOQOL-Bref scale | | | | | | |
| General QOL | 4.25 | 0.69 | 4.02 | 0.65 | 2.54 | .0 |
| General health | 4.16 | 0.79 | 3.97 | 0.86 | 1.67 | – |
| Physical | 23.25 | 3.24 | 22.00 | 3.20 | 2.91 | .004 |
| Psychological | 22.10 | 3.19 | 20.79 | 2.68 | 3.35 | .001 |
| Social | 11.54 | 2.07 | 11.14 | 1.71 | 1.56 | – |
| Environment | 30.85 | 4.39 | 29.39 | 3.89 | 2.64 | .009 |
| Total QOL | 96.15 | 10.70 | 91.31 | 9.35 | 3.61 | .0001 |
| Religiosity | | | | | | |
| General religiosity | 6.22 | 2.25 | 6.47 | 1.49 | 1.02 | – |
| Religious belief | 7.60 | 2.10 | 6.88 | 1.81 | 2.75 | .006 |

Yoon (2004) found that an individual's ▶ life satisfaction is positively related to measures of strong religious attachment in the sense of being willing to commit to attending religious services frequently. With a nationally representative sample of noninstitutionalized US adults, Maselko and Kubansky (2006) found that weekly public religious activity was significantly associated with better health and well-being. This relationship was stronger for men than women and was influenced by denominational affiliation. However, other studies have failed to find an association between religiosity and QOL, but the positive results outnumber the negative ones.

## Arab Studies on QOL and Religiosity

Based on Veenhoven's (2011) conception of happiness as an umbrella term for all that is good and its use interchangeably with well-being and QOL, there were a number of Arabic studies in this domain. Using samples from Algeria, Egypt, Kuwait, and Saudi Arabia, significant associations were found between religiosity and happiness, ▶ SWB, satisfaction with life, ▶ self-esteem, mental health (positive), and psychopathology including ▶ anxiety, depression, and neuroticism (negative) (see Abdel-Khalek, 2006, 2007, 2008, 2009a, 2010a, 2011a, c, d, e; Abdel-Khalek & Lester, 2007; Abdel-Khalek & Naceur, 2007; Tiliouine, Cummins, & Davern, 2009).

### QOL and Religiosity in a Muslim Sample

Abdel-Khalek (2011e) recruited a convenience sample of 224 Muslim, Kuwaiti, college students from several majors. Their ages ranged between 18 and 28 years. They responded to the WHOQOL-Bref scale (WHOQOL Group, 1998), along with self-rating scales to assess religiosity and religious belief, among other SWB variables. All the scales have good ▶ reliability and validity.

Results revealed that men obtained significantly higher mean scores than did their female counterparts on general QOL, physical, psychological, environmental, and the total QOL score as well as religious belief (Table 1).

Table 2 presents the correlations between the scales among the combined sample of men and women because the differences between the correlation matrices in men and women were minimal. As can be seen from Table 2, religiosity correlated significantly with four out of the six QOL subscales. Further, religious belief was significantly correlated with all QOL subscales.

## Discussion

Previous studies have considered religiosity as a salient factor influencing QOL. Over the last three decades, the studies on this domain have become an abundant area. However, the vast majority of research literature was limited in their focus on the Christian view of religion.

**Religiosity for Muslim Students in Kuwait, Quality of Life, Table 2** Pearson correlation coefficients between religiosity and QOL among men and women ($N = 224$)

| | General QOL | Health | Physical | Psychological | Social | Environment | Religiosity | Belief |
|---|---|---|---|---|---|---|---|---|
| General QOL | – | | | | | | | |
| Health | .438** | – | | | | | | |
| Physical | .203** | .243** | – | | | | | |
| Psychological | .471** | .544** | .489** | – | | | | |
| Social | .363** | .283** | .327** | .519** | – | | | |
| Environment | .529** | .309** | .360** | .523** | .435** | – | | |
| Religiosity | .130 | .114 | .176** | .145* | .197** | .207** | – | |
| Belief | .241** | .194** | .184** | .204** | .244** | .261** | .469** | – |

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

Using a Muslim Kuwaiti sample of college students, significant and positive associations were found between QOL and both religiosity and religious belief. Based on this result, the participant with high scores on religiosity and religious belief saw himself or herself as enjoying good QOL. This result is predictable on the basis of high mean score on religiosity among Kuwaiti in comparison with American students (Thorson et al., 1997). Moreover, a rational and empirical analysis of the religion of Islam reveals that the Islamic creed is a very important element in the minds and behavior of this population in general (Abdel-Khalek, 2011b).

Based on the negative association between QOL and psychopathology (Abdel-Khalek, 2010b), multiple practices are available, in Islam proper, to relieve negative emotions, including anxiety and depression, e.g., ablution, pray five times a day, reciting Qur'an, remembering Allah, call or invocation, and fasting a whole month in the year (Ramadan), among others. It is important to note that there are specific sayings (Hadith) of Prophet Mohammed. When the Muslim repeats them, they have the ability to relieve negative emotions such as ▶ worry, distress, and sadness. These sayings and practices act as a coping mechanism and a buffer against anxiety and depression. Therefore, there is a practical possibility to use religious involvement in treating anxiety and depression in Muslim clients (Abdel-Khalek, 2009b). It should be noted that this study has recruited an almost normal set of Muslim participants, and the findings of this study do not apply to religious extremists.

## References


Abdel-Khalek, A. M. (2006). Happiness, health and religiosity: Significant relations. *Mental Health, Religion and Culture, 9*, 85–97.

Abdel-Khalek, A. M. (2007). Religiosity, happiness, health and psychopathology in a probability sample of Muslim adolescents. *Mental Health, Religion and Culture, 10*, 571–583.

Abdel-Khalek, A. M. (2008). Religiosity, health and well-being among Kuwaiti personnel. *Psychological Reports, 102*, 181–184.

Abdel-Khalek, A. M. (2009a). Religiosity, subjective well-being and depression in Saudi children and adolescents. *Mental Health, Religion and Culture, 12*, 805–815.

Abdel-Khalek, A. M. (2009b, April 1–4). *The role of religion in psychotherapy in the Middle East context*. Presented at the International Congress of the World Psychiatric Association, Florence, Italy.

Abdel-Khalek, A. M. (2010a). Religiosity, subjective well-being and neuroticism. *Mental Health, Religion and Culture, 13*, 67–79.

Abdel-Khalek, A. M. (2010b, April 16–19). *Quality of life, anxiety and depression: Negative associations in Kuwaiti college students*. Presented at the 5th biennial conference of the International Society for Affective Disorders, Vancouver, Canada.

Abdel-Khalek, A. M. (2011a). Anxiety and religiosity: Negative associations among Arab adolescents and young adults. *International Journal of Psychiatry in Clinical Practice, 15*(Suppl. 2), 15.

Abdel-Khalek, A. M. (2011b). Editorial: Islam and mental health: A few speculations. *Mental Health, Religion and Culture, 14*, 87–92.

Abdel-Khalek, A. M. (2011c). Religiosity, subjective well-being, self-esteem, and anxiety among Kuwaiti

Muslim adolescents. *Mental Health, Religion and Culture, 14*, 129–140.
Abdel-Khalek, A. M. (2011d). Subjective well-being and religiosity in Egyptian college students. *Psychological Reports, 108*, 54–58.
Abdel-Khalek, A. M. (2011e). Quality of life, subjective well-being and religiosity in Muslim college students. *Quality of Life Research, 19*, 1133–1143.
Abdel-Khalek, A. M., & Lester, D. (2007). Religiosity, health and psychopathology in two cultures: Kuwait and USA. *Mental Health, Religion and Culture, 10*, 537–550.
Abdel-Khalek, A. M., & Naceur, F. (2007). Religiosity and its association with positive and negative emotions among college students from Algeria. *Mental Health, Religion and Culture, 10*, 159–170.
Allport, G. W. (1959). Religion and prejudice. *Crane Review, 2*, 1–10.
Ferriss, A. L. (2002). Religion and the quality of life. *Journal of Happiness Studies, 3*, 199–215.
Greene, K. V., & Yoon, B. J. (2004). Religiosity, economics and life satisfaction. *Review of Social Economy, 62*, 245–261.
Koenig, H. G., McCullough, M. E., & Larson, D. B. (2001). *Handbook of religion and health*. New York: Oxford University Press.
Maselko, J., & Kubansky, L. D. (2006). Gender differences in religious practices, spiritual experience and health: Results from the US General Social Survey. *Social Science and Medicine, 62*, 2848–2860.
Novak, M. (1998). The most religious century. *New York Times*, (Op-Ed.), Sunday, May 24, 1998.
Sawatzky, R., Ratner, P. A., & Chiu, L. (2005). A meta-analysis of the relationship between spirituality and quality of life. *Social Indicators Research, 72*, 153–188.
Thoresen, C. E. (1998). Spirituality, health, and science: The coming revival? In S. Roth-Roemer, S. R. Kurpius, & C. Carmin (Eds.), *The emerging role of counseling psychology in health care* (pp. 409–431). New York: Norton.
Thorson, J. A., Powell, F. C., Abdel-Khalek, A. M., & Beshai, J. A. (1997). Constructions of religiosity and death anxiety in two cultures: The United States and Kuwait. *Journal of Psychology and Theology, 25*, 374–383.
Tiliouine, H., Cummins, R. A., & Davern, M. (2009). Islamic religiosity, subjective, well-being, and health. *Mental Health, Religion and Culture, 12*, 55–74.
Veenhoven, R. (2011). Greater happiness for a greater number: Is that possible? If so how? In K. Sheldon, T. Kashdan, & M. Steger (Eds.), *Designing positive psychology: Taking stock and moving forward* (pp. 396–409). New York: Oxford University Press.
WHOQOL Group. (1995). The World Health Organization Quality of Life assessment (WHOQOL): Position paper from the World Health Organization. *Social Science and Medicine, 41*, 1403–1409.
WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine, 28*, 551–558.
Zullig, K. J., Ward, R. M., & Horn, T. (2006). The association between perceived spirituality, religiosity, and life satisfaction: The mediating role of self-rated health. *Social Indicators Research, 79*, 255–274.

# Religious Attitudes and Volunteering in the USA

Hiromi Taniguchi and Gul Aldikacti Marshall
Department of Sociology, University of Louisville, Louisville, KY, USA

## Definition

Religiosity is associated with being religious. Specifically, it refers to religious acts, beliefs, and attitudes.

## Description

### Religiosity and Formal Volunteering

Formal ▶ volunteering is typically associated with unpaid work to help or serve others through organizations. It is affected by not only socioeconomic factors such as education, occupational status, and income and social networking factors but also cultural factors such as values and beliefs. Studies have shown that religiosity plays a strong role in formal volunteering (and charitable giving; see Watt, 1991) as it generates and manages those social and cultural factors that motivate people to engage in volunteer activities (Lam, 2002; Uslaner, 2002; Wilson & Janoski, 1995; Yeung, 2004).

Although religiosity is found to be associated with volunteering and giving (Reitsma, Scheepers, & Grotenhuis, 2006), because it is a multidimensional concept (Lam, 2002), its measurement differs among scholars. Some measure it as membership to religious groups or the frequency of religious participation, more

specifically, attendance to services (Lim & McGregor, 2012; Wilson & Janoski, 1995). Those who focus on religious participation conceptualize religiosity as social capital and argue that participation expands one's social networks and subjects the person to requests for volunteering. Viewing religiosity as social capital is common in research not only in the USA but also elsewhere (Lam, 2006; Ruiter and De Graa 2006; Yeung, 2004). Furthermore, social networks allow people to develop skills (e.g., writing letters, designing and preparing flyers, scheduling meetings, and phoning) that are transferable and useful for volunteer work (Park & Smith, 2000; Verba, Schlozman, & Brady, 1995; Wilson & Musick, 1997, 1999; Wuthnow, 1999). A new study by Lim and McGregor (2012) reveals that networks involving religious organizations are so important that individuals who never or infrequently attend religious services are more likely to volunteer if they have a network of friends that attend these services regularly. Certainly, a context can operate in the opposite direction. For example, Ruiter and De Graaf (2006), in their cross-national study of 53 countries, find that the effect of religious attendance on formal volunteering is greater in secular than religious (or "devout") countries. Meanwhile, there have been criticisms toward the narrow focus on service attendance. For example, Driskell, Lyon, and Embry (2008) suggest a more comprehensive definition of religious activities by including attendance to religious services along with participation in activities such as educational programs, choir activities, missionary activities, committee meetings, witnessing, church upkeep activities, and prayer meetings.

Some scholars have moved beyond focusing simply on membership and attendance to organized activities in the public sphere (Driskell et al., 2008; Lam, 2002; Loveland, Sikkink, Myers, & Radcliff, 2005). They have looked also into the ideational level by analyzing the effects of devotion and belief (Lam, 2002; Loveland et al., 2005). Devotion is operationalized as frequency of private prayer, which "contributes to involvement focused on meeting individual needs primarily through interpersonal relations" (Loveland et al., p. 4) and also as religious reading (Lam, 2002). Examples of measurement of religious beliefs are classifying respondents by their theological beliefs (e.g., the literal truth of the Bible), beliefs in church's involvement in ▶ social justice, and beliefs in the importance of religion in their political thinking (Lam). The reason for the focus on the ideational level is that values and beliefs which are learned in a religious tradition influence one's thinking about helping others (Einolf, 2011).

Although religion is likely to promote formal volunteering, there are variations among religious traditions. The findings on differences are inconsistent, but for the most part, studies show that overall Protestants are more likely to volunteer than Catholics (Lam, 2002, 2006; Wuthnow, 1991), yet liberal Protestants (such as Presbyterians and Unitarians) are more likely to volunteer in multiple areas (secular and religious), whereas many conservative Protestants (e.g., Baptists) volunteer merely for religious activities including missionary work (Wilson & Janoski, 1995). According to Welch, Sikkink, and Lovelenad (2007), Liberal Protestants' higher rate of volunteering has to do with the higher level of ▶ trust they have toward strangers. Based on his study in the USA and Canada, Uslaner (2002) argues that conservative Christians are more likely than other Christians to devote their time to religious volunteering but not to secular volunteering. A similar result is found in Europe where Pentecostals and Jehovah's Witnesses tend to volunteer only for church-related activities while Lutherans' inclinations are to volunteer for nonchurch activities or not to volunteer at all (Yeung, 2004).

Despite the great level of attention scholars have paid to religious participation and membership as key predictors of volunteering, the effects of religious ▶ attitudes have not been explored with equal rigor. In response, Taniguchi and Thomas (2011), in their study of data from the Midlife Development in the United States (MIDUS II 2004–2006), consider multiple variables on religious attitudes along with behaviors. Their study defines religious attitude broadly and includes religiously based feelings of exclusiveness and inclusiveness and openness

to other religious teachings and examines whether and how these attitudes are associated with formal volunteering.

### Religious Exclusiveness

Social ▶ identity theory (Tajfel & Turner, 1986) suggests that when members of a group perceive themselves to be legitimately superior to others, they have a tendency to resist any attempt by members of other groups to change the intergroup (relational) status quo. Wuthnow (1999) notes the isolationist stance of the teachings of evangelical churches stressing the autonomy of the local congregation from the outside world and argues that this stance has contributed to "the ambiguities of evangelical ▶ civic engagement" (p. 359). In a similar vein, Uslaner (2001) emphasizes the tendencies of the so-called particularized trusters (those who trust only their own kind) defined by religious membership, to see religious others as a threat to their way of life and thus avoid contact with them.

It is certainly true that most, if not all, religions emphasize the importance of helping those in need. One may expect that being in contact with like-minded in-group members may enhance the pressure to conform to this near universal norm of any religion and thus facilitate collective prosocial behaviors, such as volunteering and charitable giving. Yet, contrary to their expectation, Reitsma et al. (2006) found that in many parts of Europe, the more friends one shares her or his religious view with, the less likely she or he is to engage in charitable behavior. A US study by Taniguchi and Thomas (2011) finds somewhat similar results: individuals with higher religious exclusiveness, measured as an index of five items (e.g., the level of preference to be with others who are in the same religion), are more likely to do volunteer work for their religious groups, but religious exclusiveness has no effect on secular volunteering. It thus seems that one aspect of religious attitude, exclusivist stance, is ineffective in (or even detrimental to) facilitating prosocial behavior that concerns the wider community. Religiously based feelings of exclusiveness or uniqueness can generate the feelings of exclusiveness in a broader context, and this may in turn weaken individuals' concerns for others outside their own religious communities.

### Religious Inclusiveness

Yet again, one cannot dismiss the norm of almost any religion to stress the importance of helping helpless others (Hodgkinson & Weitzman, 1990; Wilson & Musick, 1997; Wuthnow, 1991, 1994, 1999). An individual's religious faith to serve the needy and poor is one commonly noted reason for volunteering. In many religious communities, volunteering tends to be more highly valued than donation presumably because individuals can better dramatize their good deeds to live up to their religious ideals (Wuthnow, 1994). Meanwhile, formal volunteering, secular volunteering in particular, can be challenging in that, to participate, individuals are often compelled to step out of their comfort zone and interact with people who are in many ways, including religiously, different from them. They are challenged to be inclusive, that is, to be tolerant of strangers, open about differences and disagreements, and willing to try out new ideas to solve problems in hand. Scholars have thus focused on generalized trust, defined as trust we place in strangers, or generalized others, as it relates to prosocial, ▶ collective action such as volunteering (Uslaner, 2001; Wang & Graddy, 2008). While religious teachings are often alike in encouraging believers to care for others who cannot help themselves, they may significantly differ in the specifics of to "which others" to reach out. Denominational variation in the extent of participation in secular volunteering found in past research likely reflects in part how broadly (or narrowly) different religious groups define helpless others and also how favorably (or unfavorably) they perceive any "charity" activity organized outside of their religious communities.

There is not much research that explicitly considers the feelings of inclusiveness toward religious others as a predictor of formal volunteering, and existing evidence is mixed. Lam (2002) finds that in the USA, individuals who support the church's role in advocating social justice (as opposed to personal ▶ spirituality) are

no more active in secular volunteering than those who do not. On the other hand, Bekkers and Schuyt (2008) find that in the Netherlands, religiously motivated ▶ altruism promotes volunteering, especially in secular areas. Taniguchi and Thomas (2011) found that in the USA, after considering religious participation and affiliation, individuals rated higher on religious inclusiveness are more likely to engage in both religious and secular volunteering. Their measure of religious inclusiveness is based on a 7-item question about the level of agreement or disagreement with statements such as "Because of your religion or spirituality, do you try to be more tolerant of differences," ". . . to be more aware of different ways to solve problems," and "to perceive things in new ways?"

### Openness to Other Denominations

Considering religious inclusiveness as a predictor of secular volunteering as noted above is implicitly based on the assumption of clear-cut religious denominational boundaries. However, this assumption is less tenable in countries such as the USA where switching denominations has been common (Newport, 1979; Wuthnow, 1988). In these countries, individuals are free to take what they agree with from different religious traditions and create their own religion "like a patchwork quilt" (Wuthnow, 1998, p. 2). In this context, when religious attitudes are considered as predictors of volunteering, it may be fruitful to include individuals' attitudes toward exploring other religious beliefs and possibly changing their faiths as a variable or indicator related to religious inclusiveness. Especially, for secular volunteering, openness to other unfamiliar religious teachings may better motivate individuals to reach out to helpless others regardless of the beliefs of the prospective recipients of help. Such religious open-mindedness may stem from disillusionment with one's own denomination because of the social connections or moral teachings it provides (or fails to provide) (Sherkat & Wilson, 1995). This open-mindedness may then promote volunteering as a way to make up for the limitations of the current religious affiliation. Moreover, it has become increasingly common for different denominations to form alliances among themselves to deal with budgetary and membership-related constraints. This creates more opportunities for individuals affiliated with certain denominations to explore an unfamiliar denomination(s) even while participating in activities organized by religious denominations to which they belong.

Despite the high ▶ prevalence of denominational switching/religious conversion in the USA, surprisingly limited attention has been paid to whether and how religious openness predicts volunteering decisions. Batson et al. (2008) conducted an experiment with a small sample ($n = 60$) of undergraduates to study a connection between religious open-mindedness and universal compassion. They found that subjects rated high on religious open-mindedness –defined as "viewing religious doubts as positive, and readiness to face existential questions without reducing their complexity" (p. 143) – did not discriminate in who they are to help. These subjects were just as willing to help someone with a fundamentalist view, as far as it was clear to them that the help would not promote religious closed-mindedness. A Finish study (Yeung, 2004) directly considered openness to other religions as a correlate of volunteering, but found no evidence for significant association between the two variables. The lack of significant evidence is likely because Finland has a low rate of religious conversion (Barro, Hwang, & McCleary, 2010). To reiterate, denominational conversion has been common in the USA. Taniguchi and Thomas (2011) found that those who are open to other religious faiths are more likely to participate in volunteer work of both religious and secular types.

## Discussion

Focusing on religiosity as a correlate of formal volunteering has moved beyond measuring it as religious affiliation and service attendance. As research on the religiosity-volunteering link has evolved, it has become clear that other aspects of religious life often better explain commitment to volunteer work. Defining religiosity broadly, one could study how individuals use religion to reflect and form attitudes on basic concerns, such as interactions with others, which may in turn shape their decisions about

volunteering in both religious and secular areas. This review focused on one such study by Taniguchi and Thomas (2011). Their finding about the role of openness to other religious faiths has implications for religious organizations across denominations and perhaps even across major religious divisions (e.g., Christianity and Buddhism). Today, in the post-denominational era, especially in the USA, leaders of religious communities are faced with the challenges of retaining current members as well as recruiting new members. Members of religious communities are increasingly more interested in exploring other denominational (or religious) faiths and more willing to put together parts of different teachings to create their own faiths (Wuthnow, 1988). This type of openness makes it difficult for religious congregations to secure volunteers for their own needs for a long term, while making interfaith communities/organizations more attractive sites for formal volunteering.

The evolving nature of religious attitudes and identity is potentially an interesting area for researchers of formal volunteering. Whereas most existing studies in this area are cross-sectional and quantitative, exploring how religious attitudes and identity change over time and how these changes shape individuals' volunteering decisions can only be studied with a longitudinal design. An ideal longitudinal study will have a component of in-depth interviews. A promising example of this line of inquiry is found in a recent study by Einolf (2011). Using open coding of 88 in-depth interviews, he discovers six themes about religious identity and expressions and prosocial behaviors. He illuminates how individuals define morality in religious terms, incorporate religion into their identity, increase commitment to their religious beliefs over time, and manage to make "an explicit connection between religion and helping others" (p. 436).

One of the weaknesses of the research on religiosity and volunteering is the limited attention to non-Christian religions and to countries other than Canada, the USA, and some of those in Europe. Despite the fact that many studies recognize all religions as significant sources of volunteering in terms of helping others, they focus their analysis on predominantly Christian countries and distinctively Christian religious traditions. Moreover, there are not many comparative studies of different religions that shed light on how religiosity is understood in different traditions and to what extent various religions converge or diverge about their influences on volunteer activities. Future research can benefit from exploring these convergences and divergences in religious attitudes, such as inclusiveness and exclusiveness, across different religions.

## Cross-References


- ▶ Cultural Capital
- ▶ Volunteering and Young Adult Values
- ▶ Volunteering Motives in Europe
- ▶ Volunteering, Theory-Based Account


## References


Barro, R., Hwang, J., & McCleary, R. (2010). Religious conversion in 40 countries. *Journal for the Scientific Study of Religion, 49*(1), 15–36.

Batson, C. D., Denton, D. M., & Vollmecke, J. T. (2008). Quest religion, anti-fundamentalism, and limited versus universal compassion. *Journal for the Scientific Study of Religion, 47*(1), 135–145.

Bekkers, R., & Schuyt, T. (2008). And who is your neighbor? Explaining denominational differences in charitable giving and volunteering in the Netherlands. *Review of Religious Research, 50*(1), 74–96.

Driskell, R. L., Lyon, L., & Embry, E. (2008). Civic engagement and religious activities: Examining the influence of religious tradition and participation. *Sociological Spectrum, 28*(5), 578–601.

Einolf, C. (2011). The link between religion and helping others: The role of values, ideas, and language. *Sociology of Religion, 72*(4), 435–455.

Hodgkinson, V. A., & Weitzman, M. S. (1990). *Giving and volunteering in the United States: Findings from a national survey*. Washington, DC: Independent Sector.

Lam, P. (2002). As the flocks gather: How religion affects voluntary association participation. *Journal of the Scientific Study of Religion, 41*(3), 405–422.

Lam, P. (2006). Religion and civic culture: A cross-national sturdy of voluntary association membership. *Journal of the Scientific Study of Religion, 45*(2), 177–193.

Lim, C., & McGregor, C. (2012). Religion and volunteering in context: Disentangling the contextual

effects of religion on volunteer behavior. *American Sociological Review, 77*(5), 747–779.
Loveland, M., Sikkink, D., Myers, D., & Radcliff, B. (2005). Private prayer and civic involvement. *Journal of the Scientific Study of Religion, 44*(1), 1–14.
Newport, F. (1979). The religious switcher in the United States. *American Sociological Review, 44*(4), 528–552.
Park, J. Z., & Smith, C. (2000). 'To whom much has been given': Religious capital and community voluntarism among churchgoing Protestants. *Journal for the Scientific Study of Religion, 39*(3), 272–286.
Reitsma, J., Scheepers, P., & Grotenhuis, M. (2006). Dimensions of individual religiosity and charity: Cross-national effect differences in European countries? *Review of Religious Research, 47*(4), 347–362.
Ruiter, S., & De Graaf, N. (2006). National context, religiosity, and volunteering: Results from 53 countries. *American Sociological Review, 71*(2), 191–210.
Sherkat, D. E., & Wilson, J. (1995). Preferences, constraints, and choices in religious markets: An examination of religious switching and apostasy. *Social Forces, 73*(3), 993–1026.
Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago: Nelson-Hall.
Taniguchi, H., & Thomas, L. (2011). The influences of religious attitudes on volunteering. *Voluntas: International Journal of Voluntary and Nonprofit Organizations, 22*(2), 335–355.
Uslaner, E. M. (2001). Volunteering and social capital: How trust and religion shape civic participation in the United States. In P. Dekker & E. Uslaner (Eds.), *Social capital and participation in everyday life* (pp. 104–117). London: Routledge.
Uslaner, E. (2002). Religion and civic engagement in Canada and the United States. *Journal of the Scientific Study of Religion, 41*(2), 239–254.
Verba, S., Schlozman, K. L., & Brady, H. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.
Wang, L., & Graddy, E. (2008). Social capital, volunteering, and charitable giving. *Voluntas: International Journal of Voluntary and Nonprofit Organizations, 19*(1), 23–42.
Watt, D. (1991). United States: Cultural challenges to the voluntary sector. In R. Wuthnow (Ed.), *Between states and mar1cets: The voluntary sector in comparative perspective* (pp. 243–287). Princeton, NJ: Princeton University Press.
Welch, M., Sikkink, D., & Lovelenad, M. (2007). The radius of trust: Religion, social embeddedness and trust in strangers. *Social Forces, 86*(1), 23–46.
Wilson, J., & Janoski, T. (1995). The contribution of religion to volunteer work. *Sociology of Religion, 56*(2), 137–152.
Wilson, J., & Musick, M. (1997). Who cares? Toward an integrated theory of volunteer work. *American Sociological Review, 62*(5), 694–713.
Wilson, J., & Musick, M. (1999). Attachment to volunteering. *Sociological Forum, 14*(2), 243–272.
Wuthnow, R. (1988). *The restructuring of American religion*. Princeton, NJ: Princeton University Press.
Wuthnow, R. (1991). *Acts of compassion*. Princeton, NJ: Princeton University Press.
Wuthnow, R. (1994). *God and mammon in America*. New York: The Free Press.
Wuthnow, R. (1998). *After heaven: Spirituality in America since the 1950s*. Berkeley, CA: University of California Press.
Wuthnow, R. (1999). Mobilizing civic engagement: The changing impact of religious involvement. In T. Skocpol & M. P. Fiorina (Eds.), *Civic engagement in American democracy* (pp. 331–363). Washington, DC: Brookings Institution Press.
Yeung, A. (2004). An intricate triangle-religiosity, volunteering, and social capital: The European perspective, the case of Finland. *Nonprofit and Voluntary Sector Quarterly, 33*(3), 401–422.

## Religious Beliefs and Psychiatric Symptoms

Kevin J. Flannelly[1], Nava R. Silton[2] and Kathleen Galek[3]
[1]Center for Psychosocial Research, Massapequa, NY, USA
[2]Psychology, Marymount Manhattan College, New York, NY, USA
[3]HealthCare Chaplaincy, New York, NY, USA

### Synonyms

Beliefs, spiritual; Evolutionary threat assessment system (ETAS) theory; Mental disorders and religious beliefs; Psychopathology

### Definition

Religious beliefs include beliefs about God, life after death, good, and evil.

Psychiatric symptoms are defined in the various versions of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*.

## Description

### Beliefs and Psychiatric Symptoms

Cognitive therapists and theorists Aaron Beck and Paul Gilbert claim that many psychiatric disorders result from distorted beliefs about the dangerousness of the world (Flannelly, Ellison, Galek, & Koenig, 2008; Flannelly & Galek, 2010; Gilbert, 1998). These distorted and dysfunctional beliefs about the dangerousness of the world and one's personal safety are central features in general ▶ anxiety, social anxiety, panic attack, paranoia, and other psychiatric disorders (Beck, Emery, & Greenberg, 1985; Beck et al., 2001; Flannelly et al., 2012; Flannelly, Galek, Ellison, & Koenig, 2010; Gilbert, 2001; Wenzel, Sharp, Brown, Greenberg, & Beck, 2006). In the case of general anxiety disorder, the dysfunctional beliefs entail general beliefs about the world, such as any strange situation should be regarded as dangerous or it is always best to assume the worst (Beck et al., 1985; Flannelly & Galek, 2010). For individuals suffering from panic attacks, the central dysfunctional belief is that common physical and emotional experiences are catastrophic (Wenzel et al., 2006), whereas individuals with obsessive-compulsive disorder believe that everyday situations are more dangerous than they really are (Brune, 2006). For individuals with paranoid delusions, the threat is more personal, in a sense, since they believe other people want to hurt them in some way or manipulate them (Beck et al., 2001; Flannelly, Koenig, Galek, & Ellison, 2007).

The brain mechanisms underlying these kinds of beliefs presumably evolved because such beliefs have survival value. The negative bias of the beliefs associated with psychiatric symptoms has survival value because they operate on the principle that it is better to be safe than sorry. In other words, it is better to treat something as harmful (i.e., a threat) when it is not than to treat it as harmless (i.e., not a threat) when it can cause harm (Flannelly et al., 2007; Gilbert, 2001). Thus, the beliefs associated with psychiatric disorders are heuristics for making decisions about how to deal with life situations, in which the heuristic is biased toward treating situations and things as threats (Gilbert, 1998; Marks & Nesse, 1994). These heuristics have the capacity to override logic when making decisions, precisely because of their survival value (Flannelly et al., 2007; Gilbert, 1998).

From this perspective, it becomes apparent that different psychiatric disorders represent the responses of threat assessment systems in the brain to a wide array of different kinds of potential threats (Gilbert, 1998; Marks & Nesse, 1994). For instance, Marks and Nesse have argued that anxiety has survival value because it increases fitness in dangerous situations, not only with respect to one's own life and health but also with respect to vital reproductive resources, including social relationships, status, and reputation. Mild general threats produce general anxiety and its increased vigilance, which Marks and Nesse see as one end of a continuum. Panic attack is the other end of the continuum, in which the "fight or flight" response is activated, even though the circumstances do not warrant it. They consider social anxiety to be a reaction to perceived threats to social relationships, status, and reputation and various phobias to be reactions to specific threats. Some forms of paranoid ideation also involve perceived threats to social relationships (Flannelly et al., 2007).

Gilbert (2001) presents a similar analysis of the threats underlying paranoia, social anxiety, OCD, and depression. Gilbert contends that most psychiatric symptoms are produced by social threats, including general anxiety. Price and other theorists also think OCD, depression, and general anxiety are the product of social threats. However, animal research indicates that general anxiety is more akin to Freud's conception of anxiety as a "fear" of an ill-defined threat (Flannelly & Galek, 2010). Specifically, Freud thought general anxiety was a reaction to a potential source of threat of harm that was undifferentiated or ill-defined, whereas true fear was a reaction to a specific object. This view is still held today, in that anxiety is said to be a response to possible future threats of danger, whereas fear is a response to present and imminent danger (Flannelly et al., 2008).

### Brief Overview of ETAS Theory

These conceptions about the relationships among beliefs, threat assessments, and psychiatric symptoms are brought together and made explicit in

Evolutionary Threat Assessment System (ETAS) Theory (Flannelly et al., 2007, 2010; Flannelly & Galek, 2010), which has gained substantial support since it was introduced in 2007 (Flannelly et al., 2012). ETAS Theory proposes the following: (a) that specific brain structures have evolved, in part, to assess the degree to which situations and animate and inanimate objects pose a potential threat of harm; (b) that different brain structures involved in ETAS rely on different neural circuits to make their threat assessments – i.e., innate, emotional, or cognitive circuitry; (c) that many kinds of psychiatric symptoms are products of these neural threat assessments; and (d) that threat assessments and, therefore, psychiatric symptoms are directly influenced by beliefs, including religious, political, and other types of beliefs. However, since some brain structures involved in threat assessments may not be influenced by cognitive input, some classes of psychiatric symptoms may not be affected by beliefs.

### Role of Beliefs in ETAS Theory

According to ETAS Theory, beliefs modulate the sensitivity of threat assessments or modulate the threshold of what constitutes a threat. Whether beliefs modulate the sensitivity of ETAS or the threshold of what constitutes a threat, their effect is to facilitate or inhibit threat assessments. In doing so, beliefs directly increase or decrease psychiatric symptomology, respectively. For example, the belief that human nature is basically evil will increase social anxiety by facilitating threat assessments related to social relationships and situations. The belief that human nature is basically good, on the other hand, will decrease social anxiety by inhibiting threat assessments related to social relationships and situations. Beliefs about human nature would be expected to affect paranoia in the same manner. In fact, recent research has shown that believing human nature is basically evil is associated with heightened psychological distress (Ellison, Burdette, & Hill, 2009). However, no studies have examined the relationship between such beliefs and specific psychiatric symptoms.

In a similar vein, belief that God will protect you from harm should decrease general anxiety disorder and other forms of anxiety by inhibiting threat assessments about the dangerousness of the world at large. Results supporting this proposition are discussed below.

The theory also proposes that beliefs interact with one another in their net effects on psychiatric symptoms. Thus, the deleterious effects of believing that human nature is basically evil may be counteracted by the belief that God will protect you from harm. There is some evidence that beliefs do interact in yielding their net effect on psychiatric symptoms, as ETAS Theory predicts (Flannelly & Galek, 2010).

### Beliefs About God and Mental Health

Belief in an available and responsive God is a major element of Christianity and other world religions, as well. These faith traditions believe that God will be available to protect and comfort the faithful when danger threatens (Bradshaw, Ellison, & Marcum, 2010). Despite the importance of belief in religious life, relatively few studies have investigated the connection between religious beliefs and mental health (Ellison et al., 2009).

Research on beliefs about God and psychological well-being began with the work of Benson and Spilka in 1973, which examined the relationship between beliefs about God and personal ▶ self-esteem. This and subsequent studies in the United State have generally found that people who believe God is loving and caring report better psychological well-being than those who do not hold these beliefs. On the other hand, people who believe God has punished or abandoned them report poorer mental health (Flannelly et al., 2010).

The few published studies that have employed psychiatric symptoms as outcome measures have mainly examined positive beliefs about God. These studies confirm there is a salubrious relationship between belief in a close and caring God and mental health among probability samples of American adults (Bradshaw et al., 2010; Bradshaw, Ellison, & Flannelly, 2008; Flannelly et al., 2010). This research indicates that belief in a close and loving God has a salubrious association with a number of classes of psychiatric symptoms, including depression, general anxiety, social anxiety, obsession-compulsion, and paranoia (Bradshaw et al., 2008; Flannelly et al., 2010). This finding has been interpreted in light of ETAS

Theory to mean that a close and loving God reduces psychiatric symptoms because it implies that God will protect one from harm. This interpretation is supported by findings that common beliefs about God that do not imply protection have little or no salutary association with psychiatric symptoms (Flannelly et al., 2010).

### Afterlife Beliefs and Mental Health

Belief in an afterlife is a central tenet of the world's major religions, including Buddhism, Christianity, Hinduism, Islam, and to a lesser extent Judaism (Ellison et al., 2009; Flannelly et al., 2008). Belief in life after death also plays a key role in major theories of religion, some of which propose that the promise of a better life after death helps to compensate for worldly hardships and problems experienced in the present life (Bradshaw & Ellison, 2010). The promise of future spiritual rewards may also help allay fears and worries about the world (Bradshaw & Ellison, 2010; Flannelly et al., 2012). Belief in life after death is particularly high in the United States, where it appears to have increased somewhat over the years (Ellison et al., 2009; Flannelly et al., 2008, 2012).

Research investigating the connection between belief in life after death and mental health is fairly limited. However, the findings are consistent in showing that belief in an afterlife has a salutary relationship with psychiatric symptoms among national samples of American adults. A 2006 survey of the general population found that life after death was inversely related to symptoms of depression, general anxiety, social anxiety, obsession-compulsion, and paranoia (Flannelly et al., 2008, 2012). Two later studies have confirmed this salutary relationship for symptoms of anxiety and depression (Bradshaw & Ellison, 2010; Ellison et al., 2009). These later studies also indicate that belief in life after death mitigates anxiety arising from economic hardship. These research findings are consistent with ETAS Theory in demonstrating that religious beliefs are related to a variety of psychiatric symptoms.

A closer look at afterlife beliefs indicates that optimistic beliefs about life after death have salubrious associations with psychiatric symptoms, especially general anxiety, social anxiety, and paranoia (Flannelly et al., 2008). These salutary associations have been reported for belief that life after death is a life of peace and tranquility, a paradise of pleasures and delights, union with God, and reunion with family members (Flannelly et al.). These and other psychiatric outcomes have a pernicious association with pessimistic beliefs about the afterlife – i.e., that it is a pale, shadowy life or reincarnation into another life form.

Some theorists have suggested that religion – including belief in life after death – arose as means of providing a sense of security to early humans living in a dangerous world (Flannelly et al., 2008). The salutary association between optimistic beliefs about the afterlife is consistent with this premise, since they offer a better life in the future and put one's experiences in a broader context in which one's current life is only a small part of things to come. This helps to make worldly problems and concerns seem merely transitory and less threatening (Ellison et al., 2009). Given this premise, it is also understandable that pessimistic beliefs about life after death are associated with higher levels of anxiety and other psychiatric symptoms because they do not provide a sense of security in the current life and they probably raise fears about the next life (Flannelly et al., 2008).

A 2006 theoretical article by Bering proposed that belief in life after death is an element of a neurocognitive system that evolved at some point in recent human development in the context of social relations and that this system is directly involved in monitoring human social exchanges (Flannelly et al., 2012). There is evidence that such a cognitive system exists, and even evolutionary theorists of religion who doubt that belief in life after death has adaptive survival value concede that it may be a by-product of a brain system or systems that evolved to police social relationships (Flannelly et al.).

Sociologists have outlined three reasons why optimistic beliefs about life after death may reduce anxiety and other psychiatric symptoms related to social and economic inequity (Bradshaw & Ellison, 2010). First, these beliefs engender a sense of enduring, eternal significance that transcends the frustrations and failures of this life. Second, they offer an alternative perspective

by which to interpret one's earthly circumstances, including social and economic inequity. Third, optimistic beliefs about life after death may change the way one experiences daily life. Recent research suggests that optimistic afterlife beliefs may affect psychiatric symptoms by altering one's beliefs about the world at large.

A 2012 study yields some support for the contention that belief in life after death is related to beliefs about social exchanges and social ▶ equity and that optimistic afterlife beliefs may affect psychiatric symptoms by mediating other beliefs about the world (Flannelly et al., 2012). Specifically, the study found that optimistic beliefs were positively related to belief in an equitable world and negatively related to belief in a cynical world. In turn, belief in an equitable world had a salubrious association with psychiatric symptoms, whereas belief in a cynical world had a pernicious association with psychiatric symptoms, including anxiety, social anxiety, paranoia, and obsessive-compulsive disorder. From the theoretical perspective of ETAS Theory, the findings of this study suggest that religious beliefs can interact with secular beliefs about the world to influence psychiatric symptoms.

## References


Beck, A. T., Butler, A. C., Brown, G. K., Dahlsgaard, K. K., Newman, C. F., & Beck, J. S. (2001). Dysfunctional beliefs discriminate personality disorders. *Behaviour Research and Therapy, 39*(10), 1213–1225.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Bradshaw, M., & Ellison, C. G. (2010). Financial hardship and psychological distress: Exploring the buffering effects of religion. *Social Science & Medicine, 71*(1), 196–204.

Bradshaw, M., Ellison, C. G., & Flannelly, K. J. (2008). Prayer, God imagery, and symptoms of psychopathology. *Journal for the Scientific Study of Religion, 47*, 644–659.

Bradshaw, M., Ellison, C. G., & Marcum, J. P. (2010). Attachment to God, images of God, and psychological distress in a nationwide sample of Presbyterians. *International Journal for the Psychology of Religion, 20*, 130–147.

Brune, M. (2006). The evolutionary psychology of obsessive-compulsive disorder: The role of cognitive metarepresentation. *Perspectives in Biology and Medicine, 49*(3), 317–329.

Ellison, C. G., Burdette, A. M., & Hill, T. D. (2009). Blessed assurance: Religion, anxiety, and tranquility among US adults. *Social Science Research, 38*(3), 656–667.

Flannelly, K. J., Ellison, C. G., Galek, K., & Koenig, H. G. (2008). Beliefs about life-after-death, psychiatric symptomology and cognitive theories of psychopathology. *Journal of Psychology & Theology, 36*, 94–103.

Flannelly, K. J., Ellison, C. G., Galek, K., & Silton, N. R. (2012). Belief in life-after-death, beliefs about the world, and psychiatric symptoms. *Journal of Religion & Health, 51*(3), 651–662.

Flannelly, K. J., & Galek, K. (2010). Religion, evolution, and mental health: Attachment theory and ETAS theory. *Journal of Religion and Health, 49*(3), 337–350.

Flannelly, K. J., Galek, K., Ellison, C. G., & Koenig, H. G. (2010). Beliefs about God, psychiatric symptoms, and evolutionary psychiatry. *Journal of Religion & Health, 49*(2), 246–261.

Flannelly, K. J., Koenig, H. G., Galek, K., & Ellison, C. G. (2007). Beliefs, mental health, and evolutionary threat assessment systems in the brain. *Journal of Nervous and Mental Disease, 195*(12), 996–1003.

Gilbert, P. (1998). The evolved basis and adaptive functions of cognitive distortions. *British Journal of Medical Psychology, 71*(Pt 4), 447–463.

Gilbert, P. (2001). Evolutionary approaches to psychopathology: The role of natural defences. *Australian and New Zealand Journal of Psychiatry, 35*, 17–27.

Marks, I. M., & Nesse, R. M. (1994). Fear and fitness: An evolutionary analysis of anxiety disorders. *Ethology & Sociobiology, 15*(5–6), 247–261.

Wenzel, A., Sharp, I. R., Brown, G. K., Greenberg, R. L., & Beck, A. T. (2006). Dysfunctional beliefs in panic disorder: The panic belief inventory. *Behaviour Research and Therapy, 44*(6), 819–833.


# Religious Congregations and Social Justice


Nathan R. Todd, Jaclyn D. Houston and Charlynn A. Odahl
Department of Psychology, DePaul University, Chicago, IL, USA


## Definition

Religious congregations and social justice refer to the study of how religious congregations promote ▶ social justice in their local communities and society at large. Social science research to date on the topic has been conducted mainly in the fields of psychology, sociology, and ▶ anthropology.

## Description

Religious congregations are local social settings where individuals gather to practice and pursue their religion and ▶ spirituality (Chaves, 2004; Pargament, 2008; Pargament & Maton, 2000). Congregations are woven into the social fabric of society with approximately 150 million people attending over 330,000 congregations throughout the USA (Linder, 2010). There are many definitions of ▶ social justice that have been developed in political and moral philosophy based on ideas such as ▶ utilitarianism, libertarianism, and the work of philosophers such as Kant, Rawls, and ▶ Aristotle. Each definition of social justice reflects different perspectives on issues such as ▶ equality, ▶ equity, distribution of resources, and what constitutes the common good (Prilleltensky, 2012; Sandel, 2009). Recent work in philosophy and psychology frames justice in terms of ▶ capabilities (Nussbaum, 2011) or ▶ wellness as fairness (Prilleltensky, 2012). In general, definitions of ▶ social justice focus on equitable distribution of resources and the need to transform social, political, and economic systems to ensure justice for all people and groups (Prilleltensky).

Religious congregations and social justice examines how congregations promote ▶ social justice in their local communities and society at large. One example of this intersection is the integral role African American congregations played in the civil rights movement for racial equality in the United States (Calhoun-Brown, 2000). Congregational promotion of social justice has potential to promote ▶ equality and ▶ equity in society, which may contribute to other positive benefits such as increased ▶ wellness, ▶ subjective well-being, and ▶ life satisfaction. Social science research on congregations and justice moves beyond personal religious beliefs and participation to focus on the congregation as the unit of analysis to better understand the following: (a) how congregations serve as a mediating structure to connect members to work for justice (Todd & Allen, 2011), (b) congregations as empowering community settings for social justice (Maton, 2008), and (c) how congregations can aid in ▶ community capacity building and community organizing efforts (Speer et al., 2003). Each of these areas is now discussed in turn followed by a general discussion.

Individuals interface with a variety of social structures such as the family, educational systems, or religious institutions that provide a context to address social problems and to work for justice. Termed "mediating structures" in the field of sociology, these structures provide unique contexts that may influence, or "mediate," how individuals engage with society and social problems and often serve as the practical bridges between individuals and society. Todd and Allen (2011) examined congregations as mediating structures that connect individuals to social justice activities through their congregation by investigating whether or not people participated in congregational social justice activities, finding that both personal characteristics (i.e., personal religious attendance) and characteristics of the congregation (i.e., if the congregation was theologically liberal or conservative) were important in people's likelihood of engaging in congregationally based social justice activities. Qualitative research by Todd and Rufa (2012) also found that congregations may promote justice involvement by serving as both a conduit and hub for information sharing about social justice issues and ▶ volunteering opportunities both inside and outside of the congregation.

Another line of research examining ▶ civic engagement and ▶ political activities investigates how congregations may serve as mediating structures for justice by providing a context for members to develop skills for civic engagement and political engagement. Regarding civic engagement skills, scholars note that congregations may serve as natural entry points for members to hold leadership positions, to develop group speaking skills, and other organizational civic engagement skills that may promote ▶ political efficacy (Verba, Schlozman, & Brady, 1995); this may be especially true for members of ▶ marginalized communities (Schwadel, 2002). Although there is debate as to how well congregations develop civic

engagement skills for all members (Schwadel, 2002), this is potentially another way congregations serve as mediating structures, and Schwadel (2005) has found that members engaged in congregational committees are more likely also to be involved in civic committees outside of the congregation.

Encouraging engagement in ▶ political activities may be another indirect way congregations mediate social justice involvement and ▶ political empowerment. Research has examined how congregations may create this link by providing opportunities to participate in political lobbying, advocacy, and marches and how some congregations conduct voter drives and distribute voter guides (Chaves, 2004). However, based on a national random sample of over 1,200 congregations, overt political activity is peripheral to most congregations, a finding which counters the stereotypic image of congregations as deeply involved in ▶ political activities (Chaves, 2004). Even if not directly involved, Putnam and Campbell (2010) assert that congregations serve as a setting where political socialization takes place through informal friendships forged in the congregation. Although congregational political activity is the exception rather than the norm, different kinds of congregations may engage in different types of ▶ political activities. For example, Chaves (2004) found that Black Protestant congregations are more likely to engage in voter registration drives and to distribute voter guides, while White Evangelical congregations are more likely to just distribute voter guides, specifically guides from the Christian Right. Wuthnow and Evans (2002) note that Mainline Protestant congregations engage by housing political discussions and by linking members to political opportunities, but are less likely to engage in ways that influence voting. Overall, this research shows that congregations may connect members to political opportunities and facilitate ▶ political empowerment, but that congregations may do this in different ways depending on characteristics of the congregation.

Research on empowering community settings examines how congregations may mediate or empower members to engage in ▶ social justice (Maton, 2008; Maton & Salem, 1995). Maton (2008) describes empowering community settings as those that both promote the process of individual ▶ empowerment and result in a tangible increase in desired resources or a decrease in societal marginalization for people in ▶ marginalized communities. Congregations may serve as empowering community settings when they increase ▶ civic engagement or political engagement (described previously) or other social justice activities and ▶ social activism. Moreover, Maton, Domingo, and Westin (2013) describe that some congregations serve as the site for community intervention by providing physical or mental health programs, while others, about 57 % of congregations according to Chaves (2004), may serve as a community resource by providing social services that meet direct needs such as food, clothing, or shelter. Other research reports higher levels (up to 90 %) of congregations involved in social services which may be due to oversampling of larger congregations that are more likely to provide services (Chaves, 2004). Even if congregations do not directly provide social services, they may provide ▶ volunteering opportunities for such efforts (Chaves, 2004; Putnam & Campbell, 2010).

Characteristics of empowering settings have been identified such as group-based belief systems, core activities that are meaningful and congruent to setting members, an accessible and multifunctional opportunity role structure, inspirational and talented leadership, and a supportive relational environment (Maton, 2008). Research is now beginning to examine how different empowering characteristics of congregations may mediate member ▶ empowerment. For example, qualitative interviews revealed that a commitment to social justice from congregational leaders is instrumental in the congregation being involved in ▶ social justice (Todd & Rufa, 2012). More generally, differences among congregations in social service provision correspond with differences in congregational demographics and religious tradition (Chaves, 2004). Chaves notes that congregations in poor neighborhoods, those with more college graduates, and those with lower incomes provide more social services. The gender of congregational leadership also is important as Stewart-Thomas (2009) reports congregational

participation in social services is four times greater when the primary religious leader is a woman rather than a man. Furthermore, congregations that self-identify as theologically liberal report an average of two more social service programs than theologically conservative congregations (Chaves, 2004). Political involvement of congregations, as noted earlier, also depends on characteristics of congregations (Beyerlein & Chaves, 2003). Future research is needed to better understand the specific characteristics of congregations that predict different types of social justice involvement.

Congregations also may serve as empowering community settings when the congregation itself works for justice as an empowered organization and agent for ▶ social change (Maton, 2008; Maton et al., 2013). For example, congregations may lobby elected officials, sponsor marches and demonstrations, or collaborate with other organizations for social justice causes. Indeed, Ammerman (2005) notes that congregations often collaborate with community partners to achieve common goals, such as social service provision, political advocacy, or to host other neighborhood and civic groups. Furthermore, Chaves (2004) notes that 84 % of congregations engaged in social service activities did so in collaboration with an organization outside of their congregation, a finding supported by qualitative research that reveals religious leaders report collaboration as a key to community engagement (Todd & Rufa, 2012). However, certain congregations are more likely to collaborate or may collaborate with different partners. For example, congregations that are large, liberal, and educated are more likely to collaborate (Chaves, 2004). More specifically, Mainline Protestant and African American congregations are more likely to partner with secular organizations (Ammerman, 2005) or government (Chaves, 2004), whereas Conservative Protestant and Catholic congregations are more likely to partner with other congregations or religiously based organizations (Ammerman, 2005). An excellent example of congregations working for social justice is found in the role of the African American church in working for civil rights and racial ▶ equality (Calhoun-Brown, 2000). More generally, congregations may network together to promote political and ▶ social change (Speer et al., 2003; Todd, 2012). Overall, there is strong potential for religious congregations to work for ▶ social justice and ▶ social change.

## Discussion

Although the topic of religious congregations and social justice has received attention in the research literature, many questions remain unanswered. First, most research to date uses very general definitions of ▶ social justice, and future work should examine the different definitions and understandings of social justice that guide congregational activities. Todd and Rufa (2012) provide such an examination for individuals, but greater specificity at the level of the congregation is needed to understand the exact content and purpose of congregational ▶ social justice activities. As noted by many scholars, congregations have the potential to promote or inhibit social justice, depending on the understanding, definition, and application of social justice (Pargament, 2008; Pargament & Maton, 2000; Putnam & Campbell, 2010; Todd, 2011). Future research should more closely examine this tension, and the different types of social justice issues and activities congregations are pursuing under the banner of social justice. Second, most work to date focuses on Christian and Catholic congregations in the United States. Although informative, research should examine how congregations of different faiths understand and work for justice as this work may inform potential areas for collaboration among a diversity of religious traditions. Also, the connection between religious congregations and justice should be investigated in other cultural contexts beyond the United States. Overall, additional research is needed to further understand the complexity and nuance of the link between religious congregations and social justice.

## Cross-References

- ▶ Building Organizational Capacity with Community Collaboratives
- ▶ Empowerment
- ▶ Participation in Community Organizing
- ▶ Political Empowerment

- ▶ Religious Attitudes and Volunteering in the USA
- ▶ Social Change

## References


Ammerman, N. T. (2005). *Pillars of faith: American congregations and their partners*. Berkeley, CA: University of California Press.

Beyerlein, K., & Chaves, M. (2003). The political activities of religious congregations in the United States. *Journal for the Scientific Study of Religion, 42*, 229–246.

Calhoun-Brown, A. (2000). Upon this rock: The black church, nonviolence, and the civil rights movement. *PS: Political Science and Politics, 33*, 168–174.

Chaves, M. (2004). *Congregations in America*. Cambridge, MA: Harvard University Press.

Linder, E. W. (2010). *Yearbook of American and Canadian churches*. Nashville, TN: Abingdon.

Maton, K. I. (2008). Empowering community settings: Agents of individual development, community betterment, and positive social change. *American Journal of Community Psychology, 41*, 4–21.

Maton, K. I., & Salem, D. (1995). Organizational characteristics of empowering community settings: A multiple case study approach. *American Journal of Community Psychology, 23*, 631–656.

Maton, K. I., Domingo, M. R., & Westin, A. M. L. (2013). Addressing religion and psychology in communities: The congregation as intervention site, community resource, and community influence. In K. I. Pargament, A. Mahoney, & E. P. Shafranske (Eds.), *APA Handbook of Psychology, Religion, and Spirituality* (Vol. 2, pp. 613-632). Washington, DC: American Psychological Association.

Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Cambridge, MA: Harvard University Press.

Pargament, K. I. (2008). The sacred character of community life. *American Journal of Community Psychology, 41*, 22–34.

Pargament, K. I., & Maton, K. I. (2000). Religion in American life: A community psychology perspective. In J. Rappaport & E. Seidman (Eds.), *Handbook of community psychology* (pp. 495–522). New York: Klewer Academic/Plenum.

Prilleltensky, I. (2012). Wellness as fairness. *American Journal of Community Psychology, 49*, 1–21.

Putnam, R. D., & Campbell, D. E. (2010). *American grace: How religion divides and unites us*. New York: Simon & Schuster.

Sandel, M. J. (2009). *Justice: What's the right thing to do?* New York: Farrar, Straus, and Giroux.

Schwadel, P. (2002). Testing the promise of the churches: Income inequality in the opportunity to learn civic skills in Christian congregations. *Journal for the Scientific Study of Religion, 41*, 565–575.

Schwadel, P. (2005). Individual, congregational, and denominational effects on church members' civic participation. *Journal for the Scientific Study of Religion, 44*, 159–171.

Speer, P. W., Ontkush, M., Schmitt, B., Raman, P., Jackson, C., Rengert, K. M., et al. (2003). Praxis – The intentional exercise of power: Community organizing in Camden, New Jersey. *Journal of Community and Applied Social Psychology, 13*, 399–408.

Stewart-Thomas, M. (2009). Gendered congregations, gendered service: The impact of clergy gender on congregational social service participation. *Gender, Work and Organization, 17*, 406–432.

Todd, N. R. (2011). Community psychology and liberation theologies: Commonalities, collaboration, dilemmas. *Journal of Psychology and Theology, 39*, 200–210.

Todd, N. R. (2012). Religious networking organizations and social justice: An ethnographic case study. *American Journal of Community Psychology, 50*, 229–245.

Todd, N. R., & Allen, N. E. (2011). Religious congregations as mediating structures for social justice: A multilevel examination. *American Journal of Community Psychology, 48*, 222–237. doi:10.1007/s10464-010-9388-8.

Todd, N. R., & Rufa, A. K. (2012). Social justice and religious participation: A qualitative investigation of Christian perspectives. Advance Online Publication. *American Journal of Community Psychology* doi:10.1007/s10464-012-9552-4 .

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

Wuthnow, R., & Evans, J. H. (Eds.). (2002). *The quiet hand of God: Faith-based activism and the public role of mainline Protestantism*. Los Angeles: University of California Press.


## Religious Coping

- ▶ Health-Related Quality of Life and Reliance on God's Help

## Religious Doubting

- ▶ Spiritual Struggles

## Religious Freedom in Israel

- ▶ Israeli Democracy Index

## Religious Fundamentalism

▶ Democracy and Islam in the Middle East

## Religious Nonbelievers' Psychological Distress

Samuel R. Weber[1], James W. Lomax II[2] and Kenneth I. Pargament[3]
[1]Department of Psychiatry, Wexner Medical Center, The Ohio State University, Columbus, OH, USA
[2]Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA
[3]Bowling Green State University, Bowling Green, OH, USA

### Synonyms

Agnosticism; Atheism

### Definition

Religious nonbelievers are individuals who do not subscribe to religious beliefs. Nonbelievers may be subdivided into two broad categories: atheists (certain nonbelievers) and agnostics (uncertain nonbelievers). Various forms of psychological ▶ distress are experienced by religious nonbelievers, and nonbelievers have sources of distress unique to them.

### Description

Research has demonstrated that there are some areas in which nonbelievers experience greater amounts of ▶ distress compared to believers. Studies have shown that nonbelievers have greater difficulty forgiving God (Exline, Yali, & Lobel, 1999) and have more ▶ anger toward God than do believers (Exline, Park, Smith, & Carey, 2011). Although it may sound strange for those without a belief in God to demonstrate ▶ anger toward God, these nonbelievers may have a past history of religious belief; a small number of self-defined atheists in Exline's research also stated they were "unsure" about God's existence. Death ▶ anxiety associated with leaving dependents behind was greater among atheists than Catholics (James & Wells, 2002). Swedish atheists and agnostics scored higher than believers on a Death Depression Scale (which actually measures ▶ anxiety better than depression), indicating greater ▶ anxiety about death among nonbelievers (Lundh & Radon, 1998). Jews who had lost their faith described a sense of isolation, a desire to conceal their nonbelief from their community, and a sense of loneliness in their ▶ distress (Herzbrun, 1999). Among British respondents to the World Health Organization Quality of Life – Spiritual, Religious, and Personal Beliefs questionnaire – believers reported lower negative feelings and better personal relationships than agnostics. Atheists scored lower in ▶ spiritual quality of life than believers and agnostics (O'Connell & Skevington, 2010).

Other research has shown areas in which nonbelievers experience comparable or less ▶ distress than believers do. No difference was noted in depressive symptoms when comparing British atheists, agnostics, Christians, and Muslims (Baker & Cruickshank, 2010). One study of atheists demonstrated that an attitude of "indifference toward death" correlated with less ▶ anxiety about death when compared to believers (James & Wells, 2002). Although Swedish atheists and agnostics scored higher on the Death Depression Scale than did believers, the same group demonstrated no difference in ▶ anxiety levels when measured with a Stroop interference task (Lundh & Radon, 1998). Italian atheists experienced less intense negative emotions in response to the death of Pope John Paul II compared to believing Catholics (Schmidt, Sotgiu, & Tinti, 2007). British atheists have been found to cope as well as believers with the challenges of old age by compensating for losses, actively engaging in purposeful striving, and retaining positive self-regard (Wilkinson & Coleman, 2010). Jews who lost their faith were

found to be as spiritually and psychologically healthy as believing Jews, as measured by their sense of awe, ▶ purpose in life, and awareness of tragedy (Herzbrun, 1999). British atheists and agnostics affirmed their ▶ spiritual quality of life by assigning moderate importance to "▶ meaning in life" and "spiritual strength" (O'Connell & Skevington, 2010).

Conviction in one's belief system, be it atheistic or religious, correlates with psychological ▶ health. Greater conviction is associated with improved well-being. Strong atheist beliefs are comparable to strong religious beliefs when people are coping with the challenges of aging (Wilkinson & Coleman, 2010). Greater existential certainty is associated with decreased symptoms of depression (Riley, Best, & Charlton, 2005), and Christians with high levels of religious saliency have markedly fewer depressive symptoms (Baker & Cruickshank, 2010). Weaker conviction is associated with reduced well-being. Patients unsure about religion had the highest rates of drinking and the least improvement after participating in Alcoholics Anonymous (Tonigan, Miller, & Schermer, 2002). Individuals who lost their former certainty of religious belief reported poorer ▶ quality of life compared to those who retained their convictions (Brinkerhoff & Mackie, 1993). Religious adherents with low certainty of belief may be less happy than nonbelievers. 47.3 % of low-certainty religious believers were less happy when compared to atheists, 21.9 % of low-certainty believers were less happy than agnostics, and 14.4 % were less happy than those with no religious affiliation (Mochon, Norton, & Ariely, 2011). It appears that conviction in one's beliefs correlates with psychological well-being, with the strongest believers and the most convinced atheists experiencing the best ▶ health.

A potential source of psychological ▶ distress among nonbelievers is the negative perception of nonbelievers by others. College-age religious believers consistently evaluated Christians more highly than atheists, even when given examples of Christians behaving negatively and atheists behaving positively (Hunter, 1998, 2001). University students also perceived atheists as worse than Catholics in terms of moral rightness, education, social status, political views, parenting styles, ▶ life satisfaction, and control over one's life (Jenks, 1986). A large number of Americans have indicated that they would disapprove of their children marrying atheists and that atheists were "least likely to share their vision of American society" (Edgell, Gerteis, & Hartmann, 2006). These data consistently demonstrate that society overall perceives nonbelievers more negatively than believers. This negative perception is a potential source of ▶ distress among nonbelievers.

## Discussion

Although higher levels of religious belief have been linked with improvements in physical and mental ▶ health for those with a minimum level of religiousness (Koenig, King, & Carson, 2012), it does not follow that non-religiousness is necessarily detrimental to psychological functioning. In some areas, nonbelievers demonstrate more ▶ distress when compared with believers. In other ways, they experience comparable or less ▶ distress. Research on death ▶ anxiety (James & Wells, 2002; Lundh & Radon, 1998) and ▶ spiritual quality of life (Herzbrun, 1999; O'Connell & Skevington, 2010) among nonbelievers has yielded conflicting results. However, several robust points emerge from the data. There is a clear association between nonbelievers and increased ▶ anger at (Exline et al., 2011) and nonforgiveness (Exline et al., 1999) of God. Also, greater strength of conviction in one's religious (or nonreligious) worldview correlates with increased psychological ▶ health (Baker & Cruickshank, 2010; Mochon et al., 2011; Wilkinson & Coleman, 2010).

Widespread negative bias against religious nonbelievers is a potential source of psychological ▶ distress. Atheists have even been labeled as America's most problematic group, demonstrating levels of public rejection higher than that of Muslims after 9/11 (Edgell et al., 2006). Although recent decades have seen marked improvement in societal ▶ attitudes toward racial and religious diversity, atheists continue to fall short in terms of public acceptance. This nonacceptance is a potential source of ▶ distress. Promisingly, research has shown that as religious

believers perceive an increase in atheist ▶ prevalence, their ▶ prejudice against atheists decreases (Gervais, 2011).

Religious belief is a complex component of one's psychological ▶ health. Nonbelievers experience stresses and struggles unique to their situation. As a growing minority in the United States (Pew Forum, 2009), religious nonbelievers' experiences of ▶ distress must not be overlooked.

## Cross-References

- ▶ Emotional Well-Being
- ▶ Religious Beliefs and Psychiatric Symptoms
- ▶ Spirituality
- ▶ Spiritual Struggles

## References


Baker, P., & Cruickshank, J. (2010). I am happy in my faith: The influence of religious affiliation, saliency, and practice on depressive symptoms and treatment preference. *Mental Health, Religion and Culture, 12*, 339–357.

Brinkerhoff, M. B., & Mackie, M. M. (1993). Casting off the bonds of organized religion: A religious-careers approach to the study of apostasy. *Review of Religious Research, 34*, 235–258.

Edgell, P., Gerteis, J., & Hartmann, D. (2006). Atheists as "other:" Moral boundaries and cultural membership in American society. *American Sociological Review, 71*, 211–234.

Exline, J. J., Park, C. L., Smith, J. M., & Carey, M. P. (2011). Anger toward God: Social-cognitive predictors, prevalence, and links with adjustment to bereavement and cancer. *Journal of Personality and Social Psychology, 100*, 129–148.

Exline, J. J., Yali, A. M., & Lobel, M. (1999). When God disappoints. *Journal of Health Psychology, 4*, 365–379.

Gervais, W. M. (2011). Finding the faithless: Perceived atheist prevalence reduces anti-atheist prejudice. *Personality and Social Psychology Bulletin, 37*, 543–556.

Herzbrun, M. B. (1999). Loss of faith: A qualitative analysis of Jewish nonbelievers. *Counseling & Values, 43*, 129–141.

Hunter, J. A. (1998). Inter-group evaluative bias and self-esteem among Christians. *Current Research in Social Psychology, 3*, 74–87.

Hunter, J. A. (2001). Self-esteem and in-group bias among members of a religious social category. *The Journal of Social Psychology, 141*, 401–411.

James, A., & Wells, A. (2002). Death beliefs, superstitious beliefs and health anxiety. *British Journal of Clinical Psychology, 41*, 43–53.

Jenks, R. J. (1986). Perceptions of two deviant and two nondeviant groups. *The Journal of Social Psychology, 126*, 783–790.

Koenig, H. G., King, D., & Carson, V. (2012). *Handbook of religion and health* (2nd ed.). Oxford, UK: Oxford University Press.

Lundh, L., & Radon, V. (1998). Death anxiety as a function of belief in an afterlife. A comparison between a questionnaire measure and a Stroop measure of death anxiety. *Personality and Individual Differences, 25*, 487–494.

Mochon, D., Norton, M. I., & Ariely, D. (2011). Who benefits from religion? *Social Indicators Research, 101*, 1–15.

O'Connell, K. A., & Skevington, S. M. (2010). Spiritual, religious, and personal beliefs are important and distinctive to assessing quality of life in health: A comparison of theoretical models. *British Journal of Health Psychology, 15*, 729–748.

Pew Forum on Religion & Public Life. (2009). *Faith in flux: Changes in religious affiliation in the U.S* [Online]. Retrieved from 5 April, 2013 http://pewforum.org/

Riley, J., Best, S., & Charlton, B. G. (2005). Religious believers and strong atheists may both be less depressed than existentially-uncertain people. *QJM: Monthly Journal of the Association of Physicians, 98*, 840.

Schmidt, S., Sotgiu, I., & Tinti, C. (2007). The effects of religious involvement on short-term psychological reactions to the death of Pope John Paul II: A study on an Italian sample. *Social Behavior and Personality, 35*, 417–428.

Tonigan, J. S., Miller, W. R., & Schermer, C. (2002). Atheists, agnostics and alcoholics anonymous. *Journal of Studies on Alcohol, 63*, 534–541.

Wilkinson, P. J., & Coleman, P. G. (2010). Strong beliefs and coping in old age: A case-based comparison of atheism and religious faith. *Aging & Society, 30*, 337–361.


# Religiousness

- ▶ Religion/Spiritual Fulfillment, Satisfaction with
- ▶ Spirituality and Coping in Patients with Schizophrenia
- ▶ Fertility and Religion Among US Hispanics

## Rent-Geared-to-Income

▶ Housing Affordability

## Repeatability

▶ Test-Retest Reliability

## Repeated Measures Analysis

▶ Longitudinal Data Analysis

## Repetitive Motion Disorders

▶ Musculoskeletal Diseases

## Repetitive Strain Injuries (RSIs)

▶ Musculoskeletal Diseases

## Reporting Bias for Health Status

▶ Self-Report Bias in Estimating Cross-Sectional and Treatment Effects

## Reporting Bias for Treatment Effects

▶ Self-Report Bias in Estimating Cross-Sectional and Treatment Effects

## Reporting of Indices by the Press

Stephen Morse
Centre for Environmental Strategy, University of Surrey, Surrey, UK

### Synonyms

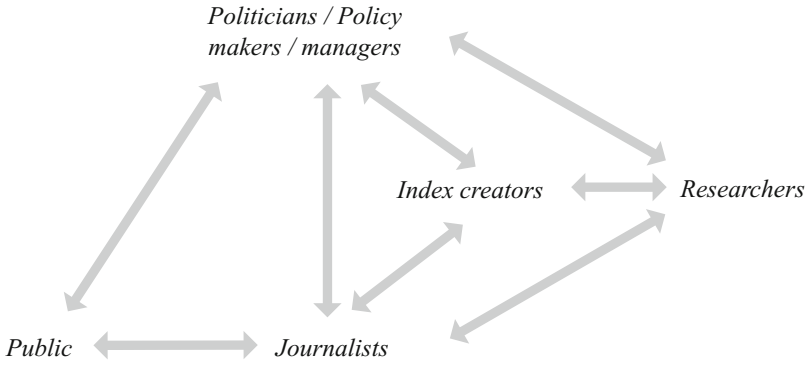
Corruption Perception Index; Ecological Footprint; Human Development Index

### Description

#### Introduction

Indices are a collection of indicators into a single value and are intended to help convey complex information to those who are meant to use it. The latter group may typically comprise nonspecialists, and indeed one of the intentions of those creating the index may be to help attract attention from the popular media (Frønes, 2007). The theory is set out in Fig. 1. Researchers generate data, often under broad research agendas laid down by policy makers, industry, and so on who control the funding. These data and the conclusions that follow from them are conveyed to policy makers via reports, etc. and are also picked up by journalists if they are deemed to be of wider relevance to their audience. Politicians are, in turn, influenced by both the public and journalists and also seek to influence these same groups. Note how many of the influence arrows in Fig. 1 have two directions. Index creators attempt to sit in the middle of these influences and use their creations to convey what they believe to be important information to policy makers and journalists. Indeed the index creators may typically be an organization, government, or nongovernment, seeking to promote a particular perspective.

However, while there is much published research on the creation of indices, and the assumptions that rest behind them, there is very little on the communication or influence of indices. This is especially so for the influence of indices on journalists and how this may be




**Reporting of Indices by the Press, Fig. 1** Simplified "influence" diagram

**Reporting of Indices by the Press, Table 1** Comparison of three indices

| Facet | Corruption Perception Index (CPI) | Human Development Index (HDI) | Ecological Footprint (EF) |
|---|---|---|---|
| Issue being captured | Corruption or more accurately "perceived corruption" | Human development. Also interpreted as a measure of quality of life and deprivation | Consumption (expressed as global productive area) |
| Units | None | None | Global hectares (gha) per capita |
| Per capita | No | No (although one of the components, GDP is expressed in terms of per capita) | Yes. Global hectares for a population (e.g., country) can be expressed on a per capita basis |
| Data source | Various corruption surveys (all based upon perception of corruption) | Three components employing data which are readily available (life expectancy, enrolment in education, GDP, population) | Requires many datasets |
| Reporting | Annual report (1990 onwards) by TI | Annual report (1995 onwards) by UNDP | Biannual report (2000 onwards) by WWF |
| Organization | Transparency international (TI; a nongovernmental organization) | United Nations Development Programme (UNDP) for the global reports | Global Footprint Network (GFN) and World Wildlife Fund (WWF) for the global reports |
| Some references | Morse (2004) | Lind (2004) | Wackernagel and Rees (1996) |



manifested in the material they present to the public.

## Reporting of Three Indices

Morse (2011) reports the results of a project designed to explore the newspaper reporting of three indices in UK national newspapers. The three indices selected for the study are:

1. Corruption Perception Index (CPI)
2. Human Development Index (HDI)
3. Ecological Footprint (EF)

Some notes on these indices are presented in Table 1. In essence all three of them were created to influence policy and behavior, and each involves the presentation of a single value for a nation-state arranged within a league table format, with best performing states at the top and worst at the bottom. The purpose of each of the indices is encapsulated in their respective titles. Hence, the CPI is intended to highlight corruption, or more precisely perceived corruption, the HDI is a broad "quality of life" index although it only has three components (education, life expectancy, and income), and the EF attempts to provide a measure of consumption (expressed in terms of global area per person rather than quantities of resource per se). The methodologies behind each of them are complex.

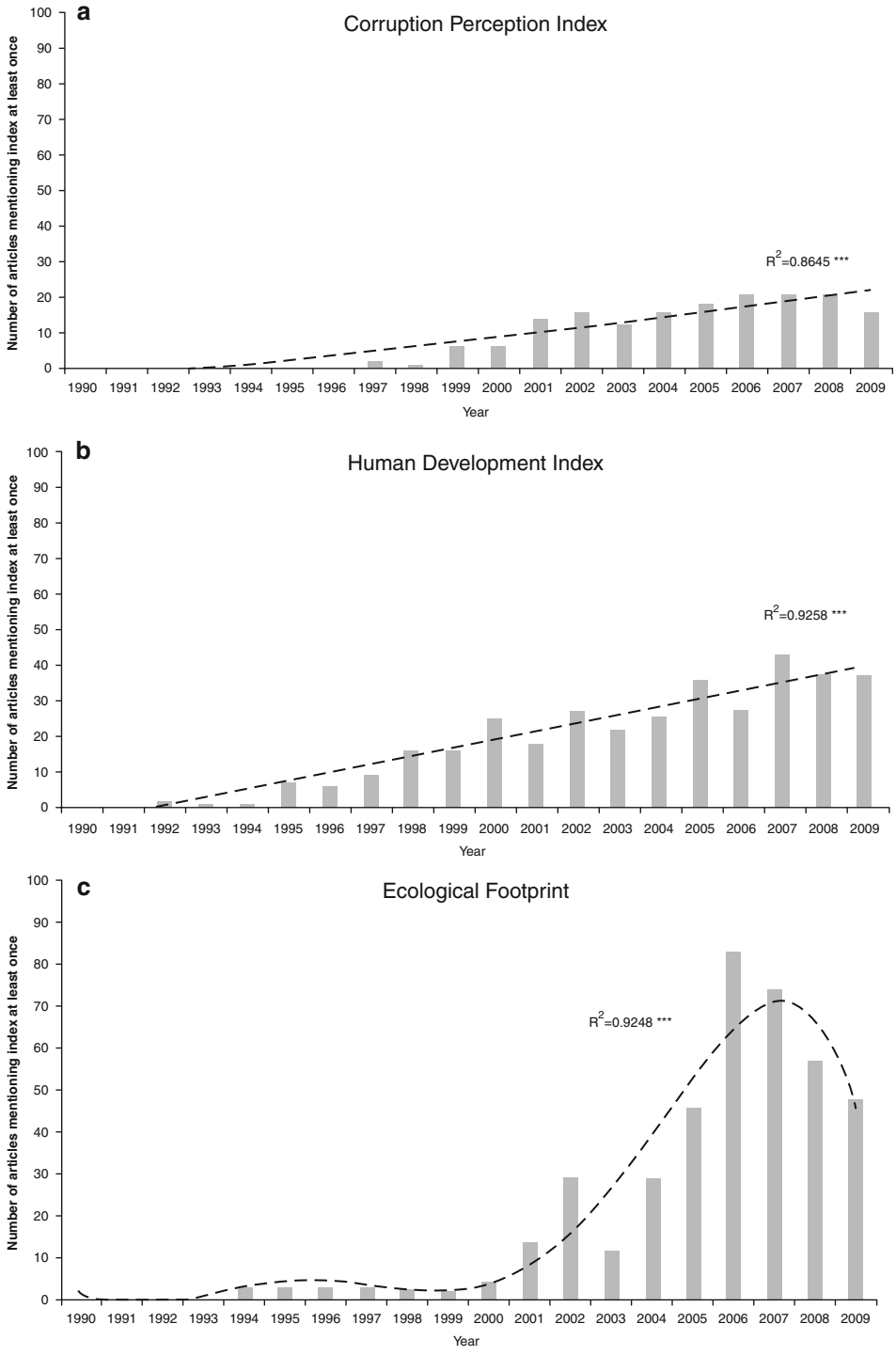
Using a database of newspaper articles for the UK, the number of articles mentioning each of these indices at least once was recorded for the period January 1990 (the first year of publication of the HDI) and December 2009 (10 years in total). The counts are shown as Fig. 2. Also shown in Fig. 2 are the results of a least squares regression fit to the data designed to illustrate the trend over time.

For the CPI and HDI, the trend over the 10 years is broadly linear. However, for the EF it is more complex, a rapid increase in articles from 2000 (the year when the EF was first adopted and publicized by the World Wildlife Fund) to a peak in 2006 (83 articles) but followed by a decline till 2009. It should be noted that although the EF was "adopted" by WWF it did exist for some years prior to that and was not a creation of the WWF. In most years, the number of articles mentioning the HDI is higher than for CPI. Over the 10 years, the total number of articles which mention the HDI at least once is 356, while for the CPI, it is 171. The number of articles mentioning the EF is higher than for the HDI and CPI, a total of 414.

How can these patterns be explained? To some extent the answer to this question is related to the ways in which the indices are promoted, but there is also an influence coming from their nature. The results suggest that reporting of the CPI is almost entirely related to the release of the Transparency International report each year. It should be noted that the CPI only appears in the Transparency International reports; no one else produces a version of the CPI, and there is no local version which compares different parts of the UK. The index is designed to provide a comparison of corruption at the level of the nation-state and as such utilizes data from a number of surveys which are pooled together to generate the CPI. It would be possible to produce a local variant of the CPI to cover places within a single country, but this would require collection of data, and the differences may be quite subtle when compared to differences between nation-states. Thus, when the TI reports are released, there is a surge in newspaper articles mentioning the index, but the interest is not sustained. The articles published around the time of the TI reports are more like announcements in terms of content, perhaps allied with a brief description of the ranking of countries as to which are best or worst and where the UK stands.

Articles mentioning the HDI also seem to be related to the publication of the Human Development Report by the UNDP, but this is not as strongly related as are the articles for the CPI and the release of the TI reports. While there are more local (intra-country or regional) versions of the HDI, this is not the case for the UK. Hence, the newspaper reporting is for the most part related to publication of the global Human Development Reports rather than anything local within the UK. There are more local versions of the HDI published for other countries, for example, the USA (Agostini & Richardson, 1997), but the government in the UK has tended to prefer other measures such as the Index of Multiple Deprivation (IMD). The lack of a local HDI limits the extent of its reporting in the UK even if important issues such as social deprivation are widely discussed and reported. Indeed it is interesting to note that the HDI is often employed by journalists to illustrate a wide range of issues surrounding poverty, quality of life, and underdevelopment but for other countries rather than the UK. For example, an article may be focused on a particular country or countries, perhaps reporting on civil strife or an aid program, and the HDI would be employed as a convenient measure to highlight their current standing. Hence, the reporting of the HDI tends to be linked to a wider range of issues than one sees with the CPI.

The EF is quite different from the CPI and HDI in the sense that while a "global" form of it is reported by the WWF, there are various other organizations which produce their own version. Regions and cities of the UK have their own form of the EF which also attracts attention from the media. Indeed the EF predates its publication in the WWF reports and was developed in the mid-1990s by William Rees and Mathis Wackernagel, then of the University of British Columbia, Canada (Wackernagel & Rees, 1996). It was originally called "appropriated carrying capacity." Indeed the idea of a "footprint"




**Reporting of Indices by the Press, Fig. 2** Number of articles mentioning the three indices at least once (January 1990 to December 2009). Figures are based upon data in Morse (2011). Also shown are a "best fit" regression line and the coefficient of determination ($R^2$) with statistical significance (*** = $P < 0.001$)

representing an impact on the planet has been picked up within wider parlance and adopted as a term in other measures, such as "carbon footprint." Thus, the term "ecological footprint" is a flexible one. It can embrace a broad notion of impact that human consumption is having on the Earth as well as an index having various forms. Nonetheless, the surge in reporting of the EF as of 2000 is no doubt due in part to the publication of the WWF reports and their form of the index. Although perhaps unsurprisingly while there is a peak of newspaper reporting of the EF following the publication of the WWF reports, akin to one sees with the TI reports and the HDR and the CPI and HDI, respectively, this is far less pronounced. The various other forms of the EF published by other groups, often at local scales in the UK, also attract the interest of journalists, and within the articles, there is also frequent use of the term "ecological footprint" to embody the idea behind the index rather than the index itself. However, the decline in reporting of the EF from 2006 is noteworthy, and one explanation may be an increasing emphasis on economic matters as concerns with the global economy gradually begin to gain momentum.

### Power of the Press

Indices are, in essence, a type of marketing tool. They are intended to help capture what can be very complex issues (e.g., corruption, human development, and the impact of consumption) and datasets into a single measure, thereby making it easier for nonspecialists to digest the message which the index is conveying. With each of the three indices discussed here, their release is within the context of an attractive report and "press pack" materials. Ultimately, of course, the hope of those creating the indices is that they will influence those who make decisions, be they the individual, household, politician, civil servant, manager, or policy maker. The publication of the CPI by Transparency International is not an end in itself. The ultimate goal is for corruption to diminish as the index helps to throw a spotlight on what is perceived to be happening across the globe. Similarly, the HDI was created by the UNDP as a means for enhancing human development. Hence, the indices are intended to be catalysts for change rather than snapshots. There are various ways of bringing about an influence, but ultimately those able to make change happen have to come under pressure, and the media are certainly one pathway for that to happen. However, just because an index is reported by the media, it does not necessarily mean that it will help bring about change. Policy is influenced by many factors, as indeed is individual behavior, and the press is just one of these; the "power of the press" is limited.

## Cross-References


- ▶ Corruption
- ▶ Ecological Footprint
- ▶ Human Development Index


## References


Agostini, S. J., & Richardson, S. J. (1997). A human development index for US cities: Methodological issues and preliminary findings. *Real Estate Economics, 25*(1), 13–41.

Frønes, I. (2007). Theorizing indicators. On indicators, signs and trends. *Social Indicators Research, 83*(1), 5–23.

Lind, N. (2004). Values reflected in the human development index. *Social Indicators Research, 66*(3), 283–293.

Morse, S. (2004). *Indices and indicators in development. An unhealthy obsession with numbers*. London: Earthscan.

Morse, S. (2011). Attracting attention for the cause. The reporting of three indices in the UK national press. *Social Indicators Research, 101*(1), 17–35.

Wackernagel, M., & Rees, W. E. (1996). *Our ecological footprint: Reducing human impact on the earth*. Gabriola Island, Canada: New Society Publishers.


# Representative Bureaucracy


- ▶ Democracy and Bureaucracy


# Representative Government


- ▶ Conceptualizing Democracy and Nondemocracy
- ▶ Measuring Democracy

# Representative Sample

Ben D'Exelle
University of East Anglia, Norwich, UK

## Definition

"Representative sampling" is a type of statistical sampling that allows us to use data from a sample to make conclusions that are representative for the population from which the sample is taken.

## Description

Quality of life researchers often only interview a sample of people instead of a whole population (a census). This has several advantages in terms of lower costs, greater depth, and the measurement of more variables. When taking a sample of a population, however, researchers need to be sure that the sample is also genuinely representative of that group.

In many cases, quality of life researchers use convenience samples, snowball samples, or quota samples. However, these do not allow the making of any statistical inference about the population from which it is taken, as they are not based on probability theory. To come to conclusions that are representative for a population, probability sampling is required, where each member of a population has known probability of being selected into a sample.

If we want to come to conclusions that are representative for a population, an assessment of the accuracy of the results calculated with the data from a sample is required, that is, how close sample values come to the true population values. However, as we usually take only one sample, we cannot make any judgment about the accuracy of a sample. That is why we make use of probability theory to estimate the likelihood that patterns in a sample hold in the population.

Every sample from the same underlying population will have different estimates. The statistic (e.g., mean or proportion) we find from a particular sample is likely to be close to the statistic of the population. The difference of the statistic between the population and the sample is called sampling error. In most of the samples, estimates come close to the value that exists in the real population, hence sampling error is limited; in only a few samples will the sample estimates be way off the mark. To assess the accuracy of a statistic, we need to have an idea of the sampling distribution of a statistic, that is, the distribution of the statistic for all possible samples from the same population of a given size. This is measured by a standard error, which is equal to the standard deviation of the sampling error (or the sampling statistic, as the statistic of the population is a constant). The larger this standard error, the lower is the precision of the sample statistic.

We could calculate a distribution of the sampling error of a statistic empirically by drawing multiple samples from the same population. However, as we only have one sample, the standard error is unknown. The standard error therefore needs to be estimated. To estimate the standard error of a sampling distribution of a statistic, we can rely on statistical theory. For example, we can make use of the central limit theorem. According to this theorem, the distribution of the mean value of a set of different samples (with equal size) from the same population approximates a normal distribution, with the average of the mean values being equal to the population average. The standard deviation of the sample means is called the standard error. As this is unknown (in reality we only have one sample) we estimate this with $\frac{\sigma}{\sqrt{n}}$, $\sigma$ being the standard deviation of the sample and $n$ the sample size. For other statistics, the formulas are usually more complicated.

Knowing the standard error and the sampling distribution, it is then common practice to calculate confidence intervals around the sample statistic for a particular confidence level. These are commonly used in the following way: For a confidence level of 95 %, we can say that we

can be 95 % sure (confident) that the statistic of the population falls within the bounds of the confidence level around the sample statistic. However, if the mean of a population is outside the 95 % confidence interval, we can say that in 95 % of cases, this would not have been observed due to sampling variation. Hence, we are 95 % confident that the underlying population is different. To calculate a confidence interval around the sample statistic, we add x times the standard error to the sample statistic to obtain the upper bound, and subtract x times the standard error to obtain the lower bound. There is a direct relation between x and the confidence level, which is commonly presented in statistical tables (for a confidence interval of 95 % around the sample mean, it can be found in a t-distribution table that x is equal to 1.96).

To increase the precision of a sample statistic, we need to reduce the standard error around the statistic. This can be achieved by increasing the sample size. According to the formula that estimates the standard error of a sample mean, for example, the standard error is inversely proportional to the square root of the sample size. This means that to reduce the standard error of a sample mean by half, we need to increase the sample size by four.

It is more common, however, for researchers to approach this question from another direction by asking what sample size is needed to make a sample representative for a particular population? For this, one needs to know the desired precision as defined by the confidence level and confidence interval, as well as the variation of the statistic in the population. As we obviously do not have any information on this variation, secondary information (e.g., from existing survey data of similar populations) is often used to obtain a realistic idea.

In addition to the precision of a sample estimate, measured in terms of sampling error which is equal to the difference between the statistic of a particular sample and the mean of the sampling distribution of that statistic, there might also be a difference between the mean of the sampling distribution of that statistic and the true population statistic, which is referred to as "bias." Such bias is caused by non-sampling error due to nonresponse (not contactable (migrated, at work, died, etc.), not interviewable (sick, language problems, too old, young, etc.), or refuses an interview), poor quality of answers, or an incomplete sample frame (e.g., out-of-date telephone list). This leads to a bias if nonrespondents, those who give poor answers or those not included in the sample frame, tend to be different from those in the sample. To assess this, it is useful to compare the results from a sample with census data or other surveys. If substantial differences exist, corrections are needed by reweighing samples.

Different methods can be used to obtain a representative sample of a population. In many cases, simple random sampling is used, where each member of a population has equal probability to be selected. A sample frame, often a complete list of all members of a population, is created, and a random subset of members is selected with the help of a randomizer. The main drawback of this method is the costs which may be prohibitively large (especially if subjects are dispersed geographically). In the latter case, another data collecting method may be preferred (e.g., telephone, internet) or another sampling design.

Multistage clustered sampling is a sampling method that is useful where geographic distance between the different subjects is large and/or where no complete sampling frame exists. In this method, the geographical area is divided in subregions (clusters), from which a sample of subregions is selected. Thereafter, for each selected subregion, a complete sampling frame is constructed, and simple random sampling is used to select a sample in each selected subregion. Additional stages may be added by subdividing selected clusters into smaller areas and randomly selecting a sample of these areas. There is an important trade-off between costs and precision. For a given sample size, a higher number of clusters (hence fewer observations in each cluster) leads to higher precision, but also higher cost. As a general rule, we need to maximize the number of clusters to sample at the initial stage as

only selecting 1 or 2 clusters at the initial stage leads to a very unrepresentative sample.

In many cases, a stratified sample or proportional sampling method produces more representative samples. This method is particularly useful when a bias is expected due to underrepresentation of certain groups because of nonresponse, poor quality of answers, or an incomplete sample frame. A variable is selected along which to stratify, for example, age, if we expect older or younger people to be under/overrepresented. Thereafter, we divide the sampling frame into separate lists for the different groups/strata (e.g., old versus young people), and we draw a random sample from each list. These samples need to be proportionate to the relative size of the various groups in the population.

Using these more complicated sample designs also affects the required sample size. For a given precision and variability, stratification reduces the required sample size (compared to simple random sampling), as sampling error due to variation between the strata is eliminated (population will be accurately represented in the sample in terms of the stratification criterion). Clustering, however, increases sample size (compared to simple random sampling) as a possible source of variation is disregarded. In particular, group level effects on the variable of interest are disregarded (e.g., some cities/villages have higher poverty levels than others; not including them in the sample increases sampling error). To calculate the required sample size of multistage, stratified and clustered samples expert advice is needed.

## Cross-References

- ▶ Multistage Sampling
- ▶ Probability Proportional Sampling
- ▶ Purposive Sampling
- ▶ Quota Sampling
- ▶ Sample Frame
- ▶ Sample Size
- ▶ Sample Survey
- ▶ Sampling Distribution
- ▶ Sampling Error
- ▶ Survey Research

## References

Bryman, A. (2008). *Social research methods* (2nd ed.). Oxford/New York: Oxford University Press, Chapters 7 & 11.

De Vaus, D. (1996). Finding a sample. In *Surveys in social research*. London: UCL press, Chapter 6.

Kish, L. (1995). *Survey sampling*. New York: John Wiley and Sons.

Levy, S., & Lemeshaw, S. (1991). *Sampling of populations: Methods and applications*. New York: John Wiley and Sons.

Poate, C. D., & Daplyn, P. F. (1993). *Data for agrarian development*. Cambridge: Cambridge University Press, Chapters 1, 2, 3,4, 7 and 8.

Treiman, D. J. (2009). Sample design and survey estimation. In *Quantitative data analysis: Doing social research to test ideas*. San Francisco: Jossey-Bass. Chapter 9.

## Reprioritization Response Shift

- ▶ Response Shift

## Reproducibility of Test Results

- ▶ Test-Retest Reliability

## Reproduction of Social Life

- ▶ Everyday Life Experience

## Reproductive Health Programs

- ▶ Family Planning

## Reproductivity of Births

- ▶ Fertility Rate

## República Mexicana

► Mexico, Quality of Life

## Rescaling

Wim Kalmijn
Erasmus Happiness Research Organization
EHERO, Erasmus University Rotterdam,
Rotterdam, The Netherlands

### Synonyms

Scale transformation

### Definition

Evaluation of QoL is done using different primary scales of measurement by different researchers. This is a problem, since it may make results of different studies incomparable. The problem can be resolved by transforming the ratings on a primary scale into ratings on a secondary scale in such a way that all secondary scales have equal lengths (for which "0" to "10" is preferred). The proposed formulation is not a fully adequate description.

Rescaling is a procedure:

- To transform the ratings of different primary scales to measure quality of life of individuals
- Into ratings on a secondary scale of measurement
- Which has a standard length for all secondary scales
- In order to make comparable different quality of life distributions
- That can be obtained by using different scales of measurement

Rescaling of primary verbal and pictorial scales includes the conversion of text or pictures into numbers that allow the application of arithmetical operations such as addition and multiplication. Since not all primary scales have equal numbers of response options, a next step is necessary to obtain secondary ratings on scales of common length.

### Description

#### Measuring Quality of Life

Many studies on quality of life (QoL) are based on surveys in which individuals in a sample from a population answer questions about their subjective well-being. In this lemma we use questions on *happiness* as an example.

#### Example: Measuring Happiness

A typical example of such a question is as follows: "Taking all things together, how would you say things are these days – would you say you are… ?" The respondent will then be asked to make a forced choice out of, e.g., four possible response options:

- "Very happy"
- "Pretty happy"
- "Not too happy"
- "Unhappy"

These four response options are elements of a *rating scale*, in this particular case, a *verbal* rating scale. The term "verbal" is used in the ► World Database of Happiness (WDH) for written textual responses. The presented response options are often referred to as *categories* and their descriptions, e.g., "pretty happy," as *labels*. The combination of a leading question, i.c., "Taking all together, how would you say things are these days?" with its complete set of $k$ response options, as presented to the respondents, is referred to as a *primary scale*. In our example $k = 4$ and this discrete scale is a verbal *four-step* or *four-point scale*.

This measurement procedure is also applicable to measuring happiness if the categories are presented as numbers (*numerical scales*) or as pictures, e.g., by using "smileys" (*pictorial scales*).

#### Different Scales of Measurement

Unfortunately, the scale in the above example is only one of the hundreds of scales that have been used in measuring happiness. A collection of all

alternatives that have ever been reported in happiness research can be found in the WDH under "Measures of Happiness" (Veenhoven, 2013).

These alternatives differ in various respects. One, there is the type and the number of response options, four in the above example, which number mostly varies between 2 and 7 for verbal and pictorial scales, whereas numerical scales may include more options, often 10 or 11. Two, the labels of the verbal response options may be different, as might be the text of the leading question; a similar problem exists in the use of pictorial scales. Finally, one has to bear in mind that translation of any text into another language usually may introduce a modification of its meaning.

As a result, much of the findings yielded with these questions are not directly comparable. This is a serious problem in QoL research, where comparison across time and place is crucial. Different strategies are used for enhancing the comparability of data yielded with different questions. Although there is a tendency to standardize measuring to the use of the 11-point numerical scale (from 0 to 10), at least to measure happiness, there is a large amount of documented research reported that is based on one of the many other original scales, so the comparison problem deserves our continuous attention to make this research accessible.

## Rescaling

Various solutions have been proposed to encounter the above problems. One class is known as "rescaling." In this approach, each response option of a $k$-step discrete measurement scale is transformed into a specific "secondary rating," which is expressed as a real number on a common interval, for which that from 0 to 10 is the usual choice. The $k$ secondary ratings together are the ratings of the secondary measurement scale. In this way a primary $k$-step scale is transformed into a secondary $k$-step scale, so *measured happiness* always remains a *discrete variable*, both before and after scale transformation. The ratings of the secondary scale are usually non-integer numbers.

Three methods of rescaling by transforming the primary ratings are:

1. ▶ Linear scale transformation, also referred to as *direct stretching* or *direct rescaling*
2. Transformation of the primary ratings into *Thurstone ratings*, ▶ Thurstone rescaling
3. Transformation into *mid-interval values*, as is proposed by Veenhoven in his Happiness scale interval study

However, the introduction of happiness in the population as a continuous ▶ Latent happiness variable offers a much more valid solution. In this approach, no transformation of the ratings of the primary measurement scale is applied. Instead, these primary ratings are included in a different conversion method of the sample measurement observations. This method is described in the Happiness, population distribution estimate lemma.

In an alternative method of measurement, respondents mark their happiness on a closed line segment, which is called a *visual analogue scale* (VAS). Responses, i.c., mark positions, are then postcoded onto a scale, which in this case is a continuous scale of measurement. Rescaling as described above is circumvented in this way.

## Statistics of the Sample Distribution on the Basis of the Secondary Scale

The $k$ ratings on the secondary scale of measurement $\{r_j;\ j = 1(1)k\}$ enable one to straightforwardly compute the sample average happiness value ($m$) as the weighted average of all $k$ different secondary ratings with the relative frequency of each rating as its weight:

$$m := \sum_{j=1}^{k} f_j r_j$$

where $f_j$ = the relative frequency of the respondents in the sample who selected the response option with secondary rating $r_j$. If, in a sample with effective size $N_e$, the response $r_j$ is selected by $n_j$ respondents, then $n_j$ is called the *absolute frequency* of the response $r_j$ and $f_j := n_j/N_e$ is its *relative frequency*. The *effective sample size* $N_e$ is the number of respondents who report one and only one of the $k$ presented response options.

The standard deviation ($s$) as a measure of the ▶ Dispersion of this frequency distribution is calculated in the standard way as

$$s := \sqrt{\sum_{j=1}^{k} f_j (r_j - m)^2}$$

In the case of a choice in favor of the ▶ linear scale transformation, the above formulae can also be used to calculate the corresponding statistics on the scale after cardinalization only, so without stretching, by simply replacing each $r_j$ with $j$.

## Cross-References

- ▶ Dispersion
- ▶ Happiness Population Distribution Parameter Estimates
- ▶ International Happiness Scale Interval Study
- ▶ Latent Happiness Variable
- ▶ Linear Scale Transformation
- ▶ Thurstone Rescaling
- ▶ World Database of Happiness

## Reference

Veenhoven, R. (2013). *Measures of happiness*, World Database of Happiness, Erasmus University Rotterdam. Available at: http://worlddatabaseofhappiness.eur.nl/hap_quer/hpi_fp.php

# Research

- ▶ Innovation Design

# Research and Development (R&D)

Sakari Kainulainen
Diaconia University of Applied Sciences, Helsinki, Finland

## Synonyms

Mode 2 knowledge production (Mode 2); Mode 2 knowledge production; Research and development and innovations (R&D&I)

## Definition

Research and development (R&D) is a broad category describing the entity of basic research, applied research, and development activities. In general research and development means systematic activities in order to increase knowledge and use of this knowledge when developing new products, processes, or services. Nowadays innovation activities are strongly tight into the concept of research and development. In the broadest meaning, research and development consists of every activity from the basic research to the (successful) marketing of a product or (effective) launching of a new process (R&D&I).

## Description

Research and development work is mostly related to business organizations. Development activities are targeted for them to development of new products and their success within the markets. A new product can be seen as the end of the chain of which starting point lies on the outputs of basic research (see Gulbrandsen).

Basic research is widely described as scientific work in order to increase understanding of fundamental principles (of specific field). In the human sciences, basic research tries to construct theories which can answer to questions like why and how phenomenon happens. On the contrary, applied research is generally carried out to solve problems of the empirical world. There are different ways to produce the knowledge and different drivers to do it. Gibbons et al. (1994) labeled traditional, investigator-oriented research as "mode 1" research and context-driven, problem-focused, and interdisciplinary knowledge production as "mode 2." The former based on theoretical interest and the latter on practical interests.

In the context of R&D, the term development is mostly linked to development of a product or a certain service. But development is a valuable term and can be used when doing intervention to some community, group, or even a nation. In this sense, it has a strong linkage to well-being, objective circumstances, or perception. The aim of the development work is not to produce a new

product for markets but more well-being for certain group of people. A good example of such work is United Nations' global work against different harasses (http://www.un.org/en/development/index.shtmlj).

In the smaller context, development can be seen as work for making life better for a certain group of people. This means solving concrete problems in everyday life in specific locations. This act comes close to applied research and especially close to one part of applied research, named action research. Action research or participatory action research (PAR) represents in a way a democratic way to do research. The role of a researcher is more as a consult than a leader of the research process. He or she is more involved to solve problems than just get new knowledge on problems. In this way, action research is part of research and development (R&D) activities (see *Action Research Journal*).

Another term used in this context is an innovation. The term innovation had firstly a commercial meaning and purpose, but nowadays it has a broader meaning covering social aspects as well. When innovation can be defined as "a commercialized product," social innovations can be defined as "making things in a different and better way in social context."

In most of the western countries, there exist public bodies whose task is to increase amount of money for research and development (R&D) activities. It is supposed to sign how dynamic and capable the society is. For example, OECD (http://www.oecd.org/development/) shares global statistics with time series on research and development activities within OECD countries and some others as well.

## Cross-References

- ▶ Action Research
- ▶ Applied Research in Quality of Life
- ▶ Canadian Research Data Centre Network
- ▶ Community-Based Participatory Research
- ▶ Development
- ▶ Economic Development
- ▶ Human Development Index (HDI)
- ▶ Index of Quality of Regional Development
- ▶ Marketing, Quality of Life
- ▶ Qualitative Indicators of Development
- ▶ Research Relationship(s)
- ▶ Social Development
- ▶ Social Indicators Research
- ▶ Survey Research
- ▶ Sustainable Development
- ▶ Sustainable Regional Development
- ▶ United Nations Commission on Sustainable Development
- ▶ United Nations Development Programme
- ▶ United Nations Research Institute for Social Development (UNRISD)
- ▶ World Bank Social Indicators of Development
- ▶ World Bank World Development Reports

## References

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Sage.

Gulbrandsen, M. The role of basic research in innovation. *Confluence*, *55*. Retrieved from http://www.cas.uio.no/Publications/Seminar/Confluence_Gulbrandsen.pdf

# Research and Development and Innovations (R&D&I)

- ▶ Research and Development (R&D)

# Research Encounters

- ▶ Research Relationship(s)

# Research Ethics

Deborah Poff
Brandon University, Brandon, MB, Canada

## Synonyms

Responsible treatment of participants in research

## Definition

Research ethics is a broad concept covering a number of principles and sets of practice in the conduct of research with respect to human beings as the subjects of or participants in research. Grounded in ▶ ethical theory and ethical principles, research ethics is premised on the ethical treatment of persons as moral agents. Research ▶ ethics also refers to the ethical treatment of animals in experimental research that uses animals as subjects of that research. While the protection and minimization of harm to animals in research is critically important, this topic focuses on the treatment of human beings.

## Description

Research ethics is not a new area of inquiry, but notably the critical inquiry about respecting human beings in research began in earnest in the mid-twentieth century following the Nuremberg trials.

As an outcome of the Nuremberg trials held after World War II for crimes against humanity which included torturous, painful, and terminal experiments with concentration camp victims and were conducted without the consent of the subjects, the Nuremberg Code was developed in 1948. The code articulates ten principles for the ethical conduct of research with human beings. All of these principles still ring true and many are contained in contemporary guidelines and frameworks for the ethical conduct of human participant research. The first and most important principle states that "the voluntary consent of the human subject is absolutely essential" (Nuremberg Code). The code further stipulates the complex nature of what it means to be in position to give voluntary consent which includes

> . . .that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved, as to enable him to make an understanding and enlightened decision. This latter element requires that, before the acceptance of an affirmative decision by the experimental subject, there should be made known to him the nature, duration, and purpose of the experiment; the method and means by which it is to be conducted; all inconveniences and hazards reasonably to be expected; and the effects upon his health or person, which may possibly come from his participation in the experiment.
>
> The duty and responsibility for ascertaining the quality of the consent rests upon each individual who initiates, directs, or engages in the experiment. It is a personal duty and responsibility which may not be delegated to another with impunity. (Nuremberg Code)

The moral imperative underlying this principle is respect and autonomy for human beings who should be treated as moral agents who are ends in themselves and not merely a means to an end in discovering new knowledge about human beings. This imperative is sometimes captured in a listing of research principles in the treatment of persons in research.

### Fundamental Principles in Research with Humans

#### Non-maleficence

This principle which is known more generally as the physician's Hippocratic Oath states, first, do no harm. The potential for doing harm is, in some cases, significant to the subjects or participants in research. In phase 1 clinical drug trials, for example, the risk of some potentially important new drug might be serious side effects which may be temporary or permanent and most seriously may result in death. In social science research, research with victims of various crimes, for example, may re-traumatize individuals who are participating in research about their experiences. In all research with human beings, the purpose, methodology, and rationale for the research must be clearly understood by the subject or participant. As well, it must be clear that consent is informed and voluntary and can be withdrawn at any time during the research. Both risks and factors which mitigate risk must also be clearly explained. In the do no harm principle, these and other key guidelines (e.g., not using a larger sample than necessary) are seen as essential to the ethical conduct of such research.

### Beneficence

More strongly than non-maleficence, the principle of beneficence is the positive principle that research should do good for participants or improve the ▶ quality of life for human beings more broadly if not for the particular individuals who participate in the research being undertaken. This implies that research with human participants must be research with a goal of improving the human condition.

Criticisms have been leveled against both researchers and practitioners that adoption of the principle of beneficence can and has led to paternalistic behaviors on the part of researchers or physicians with respect to human research subjects or patients (e.g., take this and trust me because I am the researcher or physician and know what is good for you). This in turn has led to a greater emphasis on the principle of autonomy.

### Autonomy

As listed above with respect to the first principle of the Nuremberg Code, autonomy is generally accepted as fundamental to research ethics. It is a complex principle because the prerequisites for informed consent include having free choice, an ability to understand the risks and benefits of the research, a lack of coercion or perception of coercion, and an understanding that there will be no negative consequences attached to refusing to participate in the research. Most at issue in autonomy, that is, free and informed consent is the absence of vulnerabilities.

Vulnerabilities describe a broad category of social, biological, economic, cultural, health status and cognitive conditions that may compromise free and informed consent. As such, these are complex and varied circumstances which researchers must address when ethically approaching human subject research. Children, for example, lack autonomy and the legal status to give consent so someone else, such as a parent is required to act in a child's interest if she or he is invited to participate in a research project. Socially and economically marginalized persons as well as immigrant and minority populations may lack autonomy because they do not understand fully the implications or risks of the research. Persons with mental or physical disabilities may also lack autonomy in important senses of the principle. Persons with stigmatizing conditions, like sexually transmitted diseases, or illegal occupations may feel that they must participate in research or face negative sanctions or consequences.

Many of the most egregious examples of unethical research with human subjects have been research that violated autonomy and the understanding of free and informed consent. There are also examples of unethical research where researchers ignored the fact that autonomy is a process principle and individuals must be allowed to change their mind and not continue to participate in research even if they previously agreed to do so.

### Justice

While acknowledging the issue and importance of autonomy, it is also important to recognize that persons should be allowed both to participate and not to participate in research. One of the main criticisms of much of the biological and social science research in the early and mid-twentieth century emerged during the second phase of the feminist movement. Feminist scholars argued that empirical research was frequently done solely with male subjects and then the results were extrapolated to females without any understanding of sex/gender issues. One key argument was that this practice was bad science. Another key argument was that it was unjust research.

While it is important to protect vulnerable persons, it is also important that members of vulnerable populations be given the opportunity to participate in research that will yield knowledge that may improve their quality of life. Justice requires that both the benefits and burdens of research are shared among various populations of persons.

### Summary

The recognition that human participant research requires the protection and respect for all human subjects has grown greatly in the approximately 65 years since the Nuremberg Code.

The complexity of the research ethics issues that human subject research raises mirror fundamental ▶ human rights issues more generally. The topic continues to grow as does the science and social science of the human condition.

## References


Ballantyne, A., & Rogers, W. (2007). Vulnerability and protection in research. *Harvard Health Policy Review, 8*(1), 29–41.

Bekelman, J., Li, Y., & Gross, C. (2003). Scope and impact of financial conflicts of interest in biomedical research. *Journal of the American Medical Association, 289*, 454–465.

Belmont Report. (1979). http://ohrp.osophs.dhhs.gov/humansubjects/guidance/Belmont.htm

Capron, A. (1999). Ethical and human-rights issues in research on mental disorders that may affect decision-making capacity. *The New England Journal of Medicine, 340*(18), 1430–1434.

Cho, M., Shohara, R., Schissel, A., & Rennie, D. (2000). Policies on faculty conflicts of interest at US universities. *Journal of the American Medical Association, 284*(17), 2203–2208.

Declaration of Helsinki. First developed in 1964 and revised numerous times. http://www.wma.net/policy/b3.htm

Epstein, S. (2007). *Inclusion: The politics of difference in medical research*. Chicago: University of Chicago Press.

Faden, R., Lederer, S., & Moreno, J. (1996). US medical researchers, the Nuremberg doctors trial, and the Nuremberg code. *Journal of the American Medical Association, 276*(20), 1667–1671.

Graf, C., Wager, E., Bowman, A., Flack, S., Scott-Lichter, D., & Robinson, A. (2007). Best practice guidelines on publication ethics: A publisher's perspective. *International Journal of Clinical Practice, 61* (Suppl 152), 1–26.

Hackett, E. (2002). Four observations about "Six Domains of Research Ethics". *Science and Engineering Ethics, 8*, 211–214.

Israel, M. A., & Hay, I. (2006). *Research ethics for social scientists*. Thousand Oaks, CA: Sage.

Kipnis, K. (2003). Seven vulnerabilities in the pediatric research subject. *Theoretical Medicine and Bioethic, 24*, 107–120.

Kipnis, K. (2006). Vulnerability in research subjects: A bioethical taxonomy. *Research Involving Human Participants*, 2, Online Ethics Centre for Engineering.

London, A. (2005). Justice and the international development approach to international research. *The Hastings Center Report, 35*(1), 24–37.

Nuremberg Code. http://www.ushmm.org/research/doctors/codeptx.htm

Pimple, K. (2002). Six domains of research ethics: A heuristic framework for the responsible conduct of research. *Science and Engineering Ethics, 8*, 191–205.

Poff, D. (2006). The importance of story-telling in research protocols in aboriginal communities. *Journal of Empirical Research on Human Research Ethics: An International Journal, 1*(3), 27–28.

Poff, D. (2009). Reflections on ethics in journal publications. *Journal of Academic Ethics, 7*(1–2), 51–56.

Poff, D. (2012). Academic freedom and research funding. In R. Chadwick (Ed.), *Encyclopaedia of applied ethics* (2nd ed., pp. 797–804). Netherlands: Elsevier.

Ross, L. (2004). Children in medical research: Balancing protection and access – Has the Pendulum Swung too far? *Perspectives in Biology and Medicine, 47*(4), 519–536.

See also ***JERHE*** which is a journal devoted to empirical evaluation of the efficacy of research ethics review and whether such processes and procedures work.

Shuster, E. (1997). Fifty years later: The significance of the Nuremberg code. *The New England Journal of Medicine, 13*, 1436–1440.

Sollitto, S., Hoffman, M., Lederman, R., Youngner, S., & Lederman, M. (2003). Intrinsic conflicts of interest in clinical research: A need for disclosure. *Kennedy Institute of Ethics Journal, 13*(2), 83–91.

Steneck, N. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *In Science and Engineering Ethics, 12*, 53–74.

The Encyclopedia of Applied Ethics, (20012). 2nd edition, Editor-in-Chief, Ruth Chadwick. The Netherlands: Elsevier Inc. This encyclopedia has many relevant articles on research ethics and research integrity.

The Journal of Academic Ethics publishes research ethics and research integrity articles in every issue. The Journal is published quarterly by Springer Publishing, Editor-in-Chief: Deborah C. Poff.

The Singapore Statement on Research Integrity. Singaporestatement.org.

Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (Canada), 2nd Edition.

Vollmer, S. A., & Matchett, N. (2009). Incorporating ethics in RCR courses. *Office of Research Integrity News Letter, 18*, 1.


# Research Integrity

Deborah Poff
Brandon University, Brandon, MB, Canada

## Synonyms

Responsible conduct of research

## Definition

Research integrity is a broad concept covering a number of principles and sets of practice in the conduct of research. These principles and practices are intended to ensure that researchers are ethically responsible and methodologically rigorous in the context of creating scholarly inquiry and the creation of knowledge. Research integrity is distinguished from simple error through the intentionality of the researcher to knowingly violate the ▶ norms and standards of good scholarly inquiry. Research integrity violations are not honest mistakes.

## Description

As with all research, ▶ quality of life research requires standards of research integrity. Research integrity (RI) involves a number of distinct requirements for the conduct of research that refers either to methodological rigor and standards of scientific accuracy or to the requirement for truth, honesty, and accuracy in research and in the publication of research or to the responsible treatment of research subjects and research participants by the researcher. Research integrity is sometimes referred to as the *responsible conduct of research* (RCR). Violations of RI include:

- *Fabrication or Falsification of Data or Plagiarism (FF&P)* – These violations include creating data that do not exist or falsifying the data through such activities as removing data outliers without identifying or explaining such activity or removing data that is contrary to the hypothesis being tested. Plagiarism is the intentional theft of another's ideas or work without acknowledging that this is not the researcher's work. Plagiarism can be the appropriation of published work or the theft of another research member or student of the researcher without acknowledging the real author of the research or ideas.
- *Conflict of Interest (COI)* – There are different types of conflict of interest in the context of the conduct of research. Some COI may be considered individual where a researcher or group of researchers stand to gain financially or by other direct or indirect means from certain kinds of research results, while others may be institutional, where a university or hospital may benefit. Conflicts of interest in research exist when an individual is in a position where she or he may directly or indirectly benefit from the results of the research. While the most common form of research COI is financial, other rewards include status and privilege as a consequence of enhanced reputation or profile.
- *Issues in Research Authorship* including:
  - *Ghost authorship* – where an article is written by someone else (e.g., someone employed by the funder of the research) but attributed to the researcher
  - *Gift authorship* – where a researcher is offered coauthorship without contributing anything of substance to the article
  - *Self-plagiarism* – where a researcher redundantly publishes the same material or results repeatedly or identically in numerous publications without adding to the substance of the research in any significant manner
  - *Excluding the work of others* – where a researcher excludes the authorship contribution of others who have significantly contributed to the research
- *Conducting Human Subject or Human Participant Research Without Research Ethics Approval* – Many countries now require that research undertaken in universities or teaching and research hospitals which involves human beings as subjects or participants must be reviewed by an Institutional Research Board (IRB) or ▶ Research Ethics Board (REB) before the research is undertaken. Many journals require that researchers disclose research ethics approval before considering manuscripts for publication.
- *Conducting Animal Research Without Research Ethics Approval* – Many countries now require that research undertaken in universities or teaching and research hospitals which involves animals as subjects must be reviewed by an Animal Care Committee (ACC) or Animal Research Ethics Committee before the research is undertaken.

- *Excluding Data from Research Results* (see also FF&P above) – These exclusions include:
  - *Excluding outlier data*
  - *Excluding data that disconfirm the hypotheses under question*
  - *Excluding side-effect results*

Interest in principles, guidelines, and frameworks has increased in many countries over the past 30 years as violations of research integrity and research ethics have grown in number or at least have become matters of public knowledge. In 2010, at the 2nd World Conference on Research Integrity, the Singapore Statement on Research Integrity was adopted. That document contains the following four principles:

1. Honesty in all aspects of research
2. Accountability in the conduct of research
3. Professional courtesy and fairness in working with others
4. Good stewardship of research on behalf of others

The statement also includes 14 distinct responsibilities for integrity in research, many of which are identified above. Others include:

- A commitment to rigorous peer review
- A commitment to environmental sustainability in education and research practices
- A commitment to the societal concerns of communities of study so that benefits and risks are appropriately identified in research (Singaporestatement.org)

Ultimately, RI guidelines, frameworks, and policies remind researchers that good research is grounded in integrity standards that are both methodological and moral in nature.

## References


Ballantyne, A., & Rogers, W. (2007). Vulnerability and protection in research. *Harvard Health Policy Review, 8*(1), 29–41.

Bekelman, J., Li, Y., & Gross, C. (2003). Scope and Impact of financial conflicts of interest in biomedical research. *Journal of the American Medical Association, 289*, 454–465.

Cho, M., Shohara, R., Schissel, A., & Rennie, D. (2000). Policies on faculty conflicts of interest at US Universities. *Journal of the American Medical Association, 284*(17), 2203–2208.

Epstein, S. (2007). *Inclusion: The politics of difference in medical research*. Chicago, Il: University of Chicago Press.

Graf, C., Wager, E., Bowman, A., Flack, S., Scott-Lichter, D., & Robinson, A. (2007). Best practice guidelines on publication ethics: A publisher's perspective. *International Journal of Clinical Practice, 61*(supp. 152), 1–26.

Hackett, E. (2002). Four observations about "Six Domains of Research Ethics". *Science and Engineering Ethics, 8*, 211–214.

Israel, M., & Hay, I. (2006). *Research ethics for social scientists*. Thousand Oaks, CA: Sage Publications.

Kipnis, K. (2003). Seven vulnerabilities in the pediatric research subject. *Theoretical Medicine and Bioethics, 24*, 107–120.

Kipnis, K. (2006). Vulnerability in research subjects: A bioethical taxonomy. *Research Involving Human Participants, 2*, Online Ethics Centre for Engineering.

London, A. (2005). Justice and the international development approach to international research. *The Hastings Center Report, 35*(1), 24–37.

Pimple, K. (2002). Six domains of research ethics: A heuristic framework for the responsible conduct of research. *Science and Engineering Ethics, 8*, 191–205.

Poff, D. (2006). The importance of story-telling in research protocols in aboriginal communities. *Journal of Empirical Research on Human Research Ethics: An International Journal, 1*(3), 27–28.

Poff, D. (2009). Reflections on ethics in journal publications. *Journal of Academic Ethics, 7*(1–2), 51–56.

Poff, D. (2011). Academic freedom and research funding. In R. Chadwick (Ed.), *Encyclopaedia of applied ethics* (2nd ed., pp. 797–804). Amsterdam, Netherlands: Elsevier Publishing.

Sollitto, S., Hoffman, M., Lederman, R., Youngner, S., & Lederman, M. (2003). Intrinsic conflicts of interest in clinical research: A need for disclosure. *Kennedy Institute of Ethics Journal, 13*(2), 83–91.

Steneck, N. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics, 12*, 53–74.

The Encyclopedia of Applied Ethics, (20012). 2nd edition, Editor-in-Chief, Ruth Chadwick. The Netherlands: Elsevier Inc. This encyclopedia has many relevant articles on research ethics and research integrity.

The Journal of Academic Ethics publishes research ethics and research integrity articles in every issue. The Journal is published quarterly by Springer Publishing, Editor-in-Chief: Deborah C. Poff.

The Singapore Statement on Research Integrity. Singaporestatement.org.

Vollmer, S., & Matchett, N. (2009). Incorporating ethics in RCR courses. *Office of Research Integrity News Letter, 18*, 1.

# Research Relationship(s)

Hannah Hoechner
Department of International Development, University of Oxford, Oxford, UK

## Synonyms

Context of research, social; Rapport among researchers and subjects; Reciprocity among researchers and subjects; Research encounters

## Definition

Research relationships are the connections, encounters, and exchanges between the different stakeholders in a piece of research or at a research site. These "relationships" may exist for a few hours, days or weeks, or years or even decades. While the relationship between researcher(s) and informant(s) is fundamental to all social research, the links with other individuals and institutions – such as funding bodies, gatekeepers to informants, research assistants, and the users or readership of the research findings – are equally important. These conditions influence the form and scope of research and the findings and consequences of any particular study.

## Description

The positivist assumption that social phenomena exist independent of the researcher, who is assumed to extract objective facts about society just "as a skilled surgeon can remove a kidney from a donor and use it in a different context" (Fontana and Frey 2000, p.717), has been questioned by postmodern innovations in the social sciences (see Denzin and Lincoln 2005). In interpretive and postmodern thought, the researcher is acknowledged as a human being with specific characteristics, some of which are biological (like sex, age, or color) while others are acquired (like values and basic assumptions about the world) that she/he brings with her/him to the research encounter. These characteristics – the researcher's positionality – inevitably shape "what we choose to observe, what we consider to be data, what we write about and how" (Wilson 1993, p.181). They also affect the interactions between the researcher and those being researched. How researchers present themselves to research participants (e.g., by dressing in a particular way) and how they are perceived by the latter (e.g., as a representative of a specific social group) influences the kinds of insights the researchers may subsequently gain (cf. Hallowell et al. 2005). Good research requires a constant awareness, or reflexivity, of the social context of the research encounter: "who the respondent is, who the fieldworker is, how they get on, how each thinks the other thinks, what the question means to both people, who else was there" (Lockwood 1993, p.167). This affects the data that is obtained or created and the interpretation of the research findings (cf. Finlay and Gough 2003; Denzin and Lincoln, 2000).

A careful consideration of the relationship between researcher and researched also matters for data validity and ▶ research ethics Informants are likely to reveal information only if they trust the researcher, feel at ease with the aims and modalities of the research, and the methods used enable them to express themselves and their views. Researchers will come close to understanding their informants only if they are sufficiently aware of their own social and cultural "filters" and familiar with the codes and narrative/expressive styles used by their informants (cf. Riessman 1987). It is also an ethical responsibility of the researcher to ensure that informants understand and are comfortable with the research they participate in (cf. Homan 1991; Mauthner et al. 2002).

To increase data validity and the ethical acceptability of research, particularly with vulnerable groups, participatory or action research methods have become popular in the social sciences. Many consider them to be a tool by which power differentials between researcher and researched can potentially be overcome, and the latter be empowered to have a voice in issues

concerning them. Techniques for data collection grouped under this approach include drawings, ranking exercises, group discussions, and photography. The aim of these techniques is to find a form of expression that allows research participants to convey their views and concerns freely (cf. Chambers 1997 on rapid rural appraisals; Boyden and Ennew 1997; Clark and Moss 2001 on participatory research with children). Participatory research methods have also been applied to the study of quality of life. The researchers of *Young Lives*, for instance, used activities such as body mapping and drawing of life-course timelines with children growing up in poverty to explore their subjective experiences of well-being (Crivello et al. 2009).

Yet, it is not enough just to pick a method from the "participatory menu." For research to be "truly participatory" and a positive experience for all parties involved, the methods used have to be constantly reviewed and adjusted. Ideally, research participants are involved not only in the creation of data but also in its interpretation and in the formulation of research goals as this is where power differentials are starkest (e.g., Cooke and Kothari 2001; Beazley and Ennew 2006). The burden of noncommunicable diseases among the Roma, one of the most disadvantaged ethnic groups in Europe, for example, receives far less research attention than communicable diseases which spark fears of contagion among majority populations (Hajioff and McKee 2000). Researchers in developing countries have contributed to both negative stereotypes and "incorrect but fashionable ideas" (Wilson 1993, p.180) leading to misguided policy interventions with negative results for the populations concerned. Involving, or at least consulting, research participants when interpreting data is vital for safeguarding against gross misrepresentations of the people or places being researched.

Nonetheless, handing over control of the research process to research participants is a risky undertaking that requires not only tolerance for potential "blind allies" but also sufficient time and financial resources. Few funding bodies would be willing to sponsor a research project whose topic is yet to be established! In fact, funding bodies often have not only a fairly clear research agenda but also specific views as to what constitutes acceptable methods and findings (cf. Wilson 1993, p.184). This may result in tricky situations for researchers who owe it to their informants to report "the truth" but depend on funding to continue their work.

Yet researchers' obligations toward their informants, or those they consult or discuss with, extend beyond the mere duty to represent lived realities as accurately as possible. During fieldwork it is incumbent upon researchers to ensure the protection from harm of the informants or participants of research, both during the research process and after formal data collection has ended. This encompasses the protection of informants' privacy and – where desirable – confidentiality, as well as a careful consideration of the implications that the dissemination of findings may have for their research subjects (cf. Morrow and Richards 1996).

Another big concern is that of reciprocity/compensation: What can researchers offer those who participate in their research in return for their information and the time they give out of what are undoubtedly busy days? (these concerns are particularly acute in research with materially poor respondents). As tangible benefits of academic research often fail to materialize for research participants, Wilson suggests researchers should make sure that the research process itself is rewarding for participants in terms of the human contact and intellectual stimulation it offers. Writing about ethnographic fieldwork in "developing" countries, she notes that "[l]ocals remember researchers and 'learn' from them through their personal relationships – not their monographs" (1993, p.189). At the same time, the duality of such "relationships as sources of human companionship, and as sources of data" (ibid, p.191) may well cause tensions for the researcher. What is it legitimate to collect information on and how? Should – and can – a researcher maintain an "objective" distance from her/his research subjects? Or should, on the contrary, value neutrality be dismissed, and research be "explicitly partisan – in the service of social transformation, equality, democracy,

social justice, etc." (Hammersley 2000, p.1–2)? Scheper-Hughes (1995, p.416) argues that it is the moral duty of anthropologists to take an ethical stand and develop "a political discourse on those hungry populations of the Third World that generously provide us with our livelihoods." While such positions have been criticized for being naïve, there is growing consensus that researchers have a responsibility to make their findings reach broader audiences, via, for example, community indicator development initiatives.

## Cross-References


- ▶ Privacy, an Overview
- ▶ Research Ethics


## References


Beazley, H., & Ennew, J. (2006). Participatory methods and approaches: Tackling the two tyrannies. In V. Desai & R. B. Potter (Eds.), *Doing development research* (pp. 189–199). London: Sage Publications.

Boyden, J., & Ennew, J. (1997). *Children in focus: A manual for participatory research with children*. Stockholm: Rädda Barnen.

Chambers, R. (1997). *Whose reality counts?: Putting the first last*. London: ITDG.

Clark, A., & Moss, P. (2001). *Listening to young children: The mosaic approach*. London: National Children's Bureau and Rowntree Foundation.

Cooke, B., & Kothari, U. (2001). *Participation: the new tyranny?* London: Zed Books.

Crivello, G., Camfield, L., & Woodhead, M. (2009). How Can children tell Us about their wellbeing? Exploring the potential of participatory research approaches within young lives. *Social Indicators Research, 90*, 51–72.

Denzin, N.K. & Lincoln, Y.S. (Eds.) (2005). *The SAGE Handbook of Qualitative Research* (3rd ed.). London: SAGE Publications.

Finlay, L., & Gough, B. (2003). *Reflexivity. A practical guide for researchers in health and social scienes*. Oxford: Blackwell Publishing.

Fontana, A., & Frey, J. H. (2000). The interview: From structured questions to negotiated text. In N. H. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (pp. 645–672). Thousand Oaks, CA: Sage.

Hajioff, S., & McKee, M. (2000). The health of the Roma people: A review of the published literature. *Journal of Epidemiology and Community Health, 54*, 864–869. doi:10.1136/jech.54.11.864.

Hallowell, N., Lawton, J., & Gregory, S. (2005). *Reflections on research. The realities of doing research in the social sciences*. Maidenhead: Open University Press.

Hammersley, M. (2000). *Taking sides in social research. Essays on partisanship and bias*. London: Routledge.

Homan, R. (1991). *The ethics of social research*. London: Longman.

Lockwood, M. (1993). Facts or fictions? Fieldwork relationships and the nature of data. In S. Devereux & J. Hoddinott (Eds.), *Fieldwork in developing countries* (pp. 164–178). Boulder, Colorado: Lynne Rienner Publishers.

Mauthner, M., Birch, J., Jessop, J., & Miller, T. (2002). *Ethics in qualitative research*. London: Sage.

Morrow, V., & Richards, M. (1996). The ethics of research with children. An overview. *Children and Society, 10*(2), 90–105. doi:10.1111/j.1099-0860.1996.tb00461.x.

Riessman, C. (1987). When gender is not enough: Women interviewing women. *Gender and Society, 1*(2), 172–207. doi:10.1177/089124328700100200 4.

Scheper-Hughes, N. (1995). The primacy of the ethical: Propositions for a militant anthropology. *Current Anthropology, 36*(3), 409–420.

Wilson, K. (1993). Thinking about the ethics of fieldwork. In S. Devereux & J. Hoddinott (Eds.), *Fieldwork in developing countries* (pp. 179–199). Boulder, Colorado: Lynne Rienner Publishers.


## Research Utilization


- ▶ Knowledge Transfer and Exchange


## Reserves


- ▶ Parks and Quality of Life


## Resettlement


- ▶ Migration, an Overview


## Residential Community


- ▶ Gated Communities

## Residential Environmental Satisfaction

▶ Residential Satisfaction

## Residential Mobility

▶ Housing and Aging

## Residential Satisfaction

Kathryn Terzano
Westfield State University, Westfield, MA, USA

### Synonyms

Residential environmental satisfaction

### Definition

Residential satisfaction refers to people's level of approval with conditions in their residential neighborhood.

### Description

Experts differ on whether they conceive satisfaction to be a cognitive evaluation of the affordances provided by the environment, an affective evaluation of how well the environment meets the needs of the residents, or some combination of the two kinds of evaluations with a further attitudinal component (Canter & Ress, 1982; Francescato, Weidemann, & Anderson, 1989; Galster, 1987).

This broad term includes satisfaction with the social and economic factors of the residents as well as with the built and natural environments (Mesch & Manor, 1998). The scale ranges from an individual's housing unit to the neighborhood in which that housing unit is located and even the larger surrounding area and the city. Factors include both objective ones (e.g., crime rates) and subjective ones (e.g., perceptions of crime levels and safety). Residential satisfaction is related to one's quality of life and sense of community (Hur & Morrow-Jones, 2008) and desire to move (Mudege & Zulu, 2011; Newman & Duncan, 1979).

Long-time residents and new residents differ in what they report to be the most salient factors influencing satisfaction. Those residents who have lived somewhere for a long time report that stress factors, such as noise, crime levels, and race relations, are the most important factors influencing their residential satisfaction; however, new residents cite physical appearance as being most important to them (Potter & Cantarero, 2006). Additionally, greater length of residence is positively associated with neighborhood satisfaction as long-time residents tend to have more social ties and stronger place attachment (Galster, 1987). Logically, some portion of residents who are unsatisfied with their housing or neighborhood will move away from the neighborhood, whereas more satisfied residents are likely to remain.

Certain objective factors have been well studied, including variables such as income and education levels, home owner status (home owner versus renter), and the presence of children, and these factors have been found to significantly correlate with residential satisfaction (Amerigo & Aragones, 1990). Increased traffic and noise, smells, and pollution negatively affect residential satisfaction (Hur & Morrow-Jones, 2008; Kroesen et al., 2010; Van Dyck, Cardon, Deforche, & De Bourdeaujuij, 2011). Objective factors are diverse. Other examples of objective attributes of the residential environment include housing vacancy rates, availability of grocery stores, and amount of parkland (Marans & Stimson, 2011).

Still other factors consist of aspects related to housing, street, and neighborhood design. Neotraditional designs seek to increase satisfaction and neighborliness through small setbacks, the use of front porches, the positioning of garages to the rear of houses, and narrow streets (Marans & Rodgers, 1975). However, other studies have shown that people prefer postwar suburbs to neotraditional neighborhoods

(Lovejoy, Handy & Mokhtarian, 2010; Morrow-Jones, Irwin, & Roe, 2004; Nasar & Julian, 1995). Residents of older housing units and neighborhood report greater residential satisfaction than residents of newly built neighborhoods, possibly because this is related to the strength of community ties and networks in the neighborhood (Kroesen et al., 2010; McHugh, Gober, & Reid, 1990).

Researchers have considered these factors in conceptual models used to study residential satisfaction. Marans and Rodgers (1975) presented a model of the relationships between residential satisfaction (at the scales of housing/dwelling, neighborhood, city/town, and country) and quality of life. Their model showed how the quality of place is experienced subjectively by individuals, reflecting each person's perceptions, assessments, and past experiences. In this model, the objective factors of a place are insufficient in determining its quality of residents' satisfaction with it. Francescato, Weidemann, and Anderson (1989) proposed a six-domain model of residential satisfaction based on the idea that satisfaction is largely attitudinal. Their six domains include objective environmental attitudes, individual characteristics, behavioral and normative beliefs, perceptions, emotions, and behavioral intentions.

Research on residential satisfaction is ongoing. New directions for research in this area include seeking to identify the qualities of neighborhoods that make them more satisfying and sustainable over time.

## Cross-References

- ▶ Community Satisfaction
- ▶ Housing and Aging
- ▶ Residential Mobility
- ▶ Sense of Community

## References


Amerigo, M., & Aragones, J. I. (1990). Residential satisfaction in council housing. *Journal of Environmental Psychology, 10*(4), 313–325.

Canter, D., & Ress, K. A. (1982). Multivariate model of housing satisfaction. *International Review of Applied Psychology, 32*, 185–208.

Francescato, G., Weidemann, S., & Anderson, J. R. (1989). Evaluating the built environment from the users point of view: An attitudinal model of residential satisfaction. In W. F. E. Preiser (Ed.), *Building evaluation* (pp. 181–198). New York: Plenum Press.

Galster, G. (1987). *Homeowners and neighborhood reinvestment decisions*. Durham, NC: Duke University Press.

Hur, M., & Morrow-Jones, H. (2008). Factors that influence residents' satisfaction with neighborhoods. *Environment and Behavior, 40*(5), 619–635.

Kroesen, M., Molin, E. J. E., Miedema, H. M. E., Vos, H., Janssen, S. A., & Van Wee, B. (2010). Estimation of the effects of aircraft noise on residential satisfaction. *Transportation Research Part D, 15*, 144–153.

Lovejoy, K., Handy, S., & Mokhtarian, P. (2010). Neighborhood satisfaction in suburban versus traditional environments: An evaluation of contributing characteristics in eight California neighborhoods. *Landscape and Urban Planning, 97*, 37–48.

Marans, R. W., & Rodgers, W. (1975). Toward an understanding of community satisfaction. In A. H. Hawley & V. P. Rock (Eds.), *Metropolitan America in contemporary perspective* (pp. 299–352). New York: Sage.

Marans, R. W., & Stimson, R. (2011). *Investigating quality of urban life: Theory, methods and empirical research*. New York: Springer.

McHugh, K. E., Gober, P., & Reid, N. (1990). Determinants of short- and long-term mobility expectations for home owners and renters. *Demography, 27*(1), 81–95.

Mesch, G. S., & Manor, O. (1998). Social ties, environmental perception, and local attachment. *Environment and Behavior, 30*, 504–519.

Morrow-Jones, H., Irwin, E. G., & Roe, B. (2004). Consumer preference for neotraditional neighborhood characteristics. *Housing Policy Debate, 15*, 171–202.

Mudege, N., & Zulu, E. (2011). In their own words: Assessment of satisfaction with residential location among migrants in Nairobi slums. *Journal of Urban Health, 88*, 219–234.

Nasar, J. L., & Julian, D. A. (1995). The psychological sense of community in the neighborhood. *Journal of the American Planning Association, 61*, 178–184.

Newman, S. J., & Duncan, G. J. (1979). Residential problems, dissatisfaction, and mobility. *Journal of the American Planning Association, 45*, 154–166.

Potter, J., & Cantarero, R. (2006). How does increasing population and diversity affect resident satisfaction? A small community case study. *Environment and Behavior, 38*, 605–625.

Van Dyck, D., Cardon, G., Deforche, B., & De Bourdeaujuij, I. (2011). Do adults like living in high-walkable neighborhoods? Associations of walkability parameters with neighborhood satisfaction and possible mediators. *Health & Place, 17*, 971–977.

## Residuals, Analysis of

Nand Kishor
Department of ECPS, University of British Columbia, Vancouver, BC, Canada

### Synonyms

Investigating linear regression assumptions; Robustness

### Definition

Residual analysis involves an examination of the assumptions of the statistical model fitted to the data. Residuals are the differences between the observed responses and the responses predicted by the fitted model.

### Description

Social scientists often fit linear models, characterized as ANOVA or ▶ regression, to examine their theories (Cohen, Cohen, West, & Aiken, 2003). The differences between the observed responses and responses predicted by the fitted model are called "residuals." The adequacy of the models is based on certain assumptions about these residuals. Hence, an analysis of these residuals plays an important role in examining the fit of the statistical model.

The most common and the general form of the statistical model used in the quality of life research is linear regression. Key assumptions underlying the regression model include the statistical independence and normality of residuals, linearity of relationship, constant error variance, and independence of residuals. ▶ ANOVA is a restricted case of regression where the predictors are only qualitative, but the same assumptions apply. The way to verify these assumptions is through an analysis of the residuals. When the assumptions are met, the least squares estimates in the model have desirable properties for statistical inference.

Much of the residual analysis is based on examination of graphical plots of residuals, usually the standardized or the Studentized residuals, although there are some objective tests and markers as well. By examining the pattern of residual plots, we not only spot any violation of the assumptions, but also determine the presence of ▶ outliers and influential cases, omitted variable(s), and the overall appropriateness of the model. Most statistical software have options for creating residual plots as well as the option to save residuals for further statistical analysis.

*Linearity:* A bivariate display of the outcome variable versus the predictors is often used to examine the linearity of the relationship. However, in multiple regression the interest is in the "partial" relationship between the outcome variable and a predictor controlling for all other predictors in the model. Therefore, a preferable method is an examination a scatter plot of Studentized residuals against each predictor. This "partial residual plot" is available in most software and is well-suited for displaying non-linearity. A non-random scatter will reveal non-linearity. Smoothers such as "lowess" are available in many software packages as well that can be used to detect hidden monotone non-linearity.

*Independence:* It is the residuals and not the observed values of the outcome variable that are assumed to be independent (Cook & Weisberg, 1999). In a scatter plot of standardized residuals against any time variables or any clustering of cases, a non-random pattern would suggest lack of independence. The Durbin-Watson statistic, available in most software, indicate serial correlation in the residuals. The Durbin-Watson statistic lies between 0 and 4. A value of 2 means there is no serial correlation in the residuals. Values less than 2 indicate positive serial correlation and values greater than 2 indicate negative serial correlation. Non-independence of residuals affect the standard errors and hence the significance tests.

*Constant variance:* Often referred to as "homoscedasticity," the variability of residuals around the regression line is expected to be constant regardless of the values of the outcome variable in multiple regression. If the variance is

related to values of the outcome variable, "heteroscedasticity" is present and the assumption is violated, which will affect the standard errors and hence the significance tests. A scatter plot of Studentized residuals against predicted values will show a random symmetrical scatter around the line at zero when there is homoscedasticity. In the case of heteroscedasticity, the scatter plot will reveal a funnel or a megaphone shape, or other non-random pattern, which may imply an omitted important predictor from the model. Formal tests for homoscedasticity exist as well (e.g. White, 1980) which can be computed from saved residuals.

*Normality:* Residuals can be saved and their distribution examined with Kolmogorov-Smirnov test or Shapiro-Wilk's tests of normality available in most statistical software. However, a normal probability plot or Q-Q plot of the standardized residuals can reveal the skewness and the tails of the distribution guiding remedial choices, if needed. The actual data are ranked, an expected normal value is computed, and compared with an actual normal value for each case. Points lining up along the straight diagonal that goes from lower left to upper right indicate normality. Departure from the expected straight line, particularly "snaking" signals non-normality. Non-normality is not likely to have serious consequences for the significance tests unless the sample is small, and it is in small data sets, the Q-Q plot could be misleading.

In summary, residual analysis reveals the goodness of fit of the model for statistical inference. Graphical analysis methods are the most popular because they have the advantage of illustrating the complex interrelationships which may aid decisions about model modification. It is an extremely valuable exercise for building a more stable mathematical model. A stable model is more replicable.

## Cross-References

- ▶ Data Analysis
- ▶ Outliers
- ▶ Parametric Analysis
- ▶ Robust Statistical Tests

## References

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Earlbaum Cook.

Cook, R. D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. New York: Wiley.

White, R. J. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test of heteroscedasticity. *Econometrica, 48*, 817–838.

# Resilience

Chelsea Mitamura[1], Lillian Reuman[2] and Michele Tugade[3]
[1]Vassar College, University of Wisconsin, Madison, WI, USA
[2]Vassar College, Massachusetts General Hospital, Boston
[3]Department of Psychology, Vassar College, New York, NY, USA

## Synonyms

Adaptability; Hardiness; Management of stress; Recovery; Resourcefulness;

## Definition

The term *resilience* alludes to one's ability to successfully adapt and endure under adverse circumstances as well as efficiently recover from subsequent harmful effects (e.g., Bonanno, 2004; Masten, 2007; Skodol, 2010). This is achieved by productively managing negative emotions and by practicing behavioral responses that improve the stressful situation (Lyubomirsky & Della Porta, 2010; Tugade, 2011). Some scholars conceptualize resilience as a *state* or *process* that considers context and the availability of resources and assets. Other scholars theorize resilience as a personality *trait* or focus on the developmental nature of the construct (Masten & Reed, 2002). Individuals high in trait resilience have demonstrated positive progressive outcomes despite high risk, sustained competence under

stress, and recovery from trauma. Past research suggests that certain protective factors within an individual may predict resilience in the face of stress. Such factors include positive emotionality, optimism, effective problem solving, faith, sense of meaning, ▶ self-efficacy, flexibility, impulse control, empathy, close relationships, and ▶ spirituality, among others (Bonanno, 2004; Masten & Reed, 2002; Tugade, 2011; Tugade & Fredrickson, 2004). In this way, resilient individuals typically demonstrate a strong, well-differentiated, cohesive sense of self and enduring, secure interpersonal relationships (Skodol, 2010).

## Description

Although an abundance of literature focuses on the demonstrative characteristics of resilient individuals, research on the mechanisms or processes by which resilience can be achieved is relatively understudied. One useful approach is to investigate the multiple pathways to resilience (Bonanno, 2004). For all intents and purposes, four unique pathways have been identified. These four pathways are differentiated by the quantity, intensity, and timing of stressors faced during childhood, adolescence, and adulthood.

One pathway characterizes individuals who appear relatively unscathed. These individuals may profit from positive self-evaluations and optimistic outlooks on life by adopting "rose-colored glasses" in the face of stress. Such positive appraisals are particularly useful for helping people overcome major stressful events or traumas, permitting them to bounce back from adversity and maintain ▶ homeostasis in spite of stressful experiences (Arce et al., 2009; Tugade & Fredrickson, 2004).

A second pathway characterizes individuals who, despite experiencing minor setbacks, or ▶ daily hassles, can function effectively. These stressors may include temporary academic difficulty, personal injury, and changing friendships (DeLongis, Coyne, Dakof, Folkman, & Lazarus, 1982). Facing mild stressors, these individuals develop personally effective strategies for coping and remain well-adjusted while enduring relative struggle.

A third category of resilience characterizes individuals who exhibit effective coping in spite of chronic stressors such as ▶ poverty, living with a depressed parent or spouse, or constant relocation as part of a military family. These individuals develop resilience by monitoring their emotions/resources and establishing a perception of control. As an example, the US army provides combat soldiers with resilience training (e.g., adopting optimistic perspectives), giving soldiers conceptual tools to aid them in coping with anticipated stressful circumstances (Reivich, Seligman, & McBride, 2011).

A final category of resilience characterizes individuals who have faced one or more traumatic experiences (e.g., death of a loved one, loss of a job, violent assault or natural disaster). These individuals may adapt effectively from a host of stressful experiences by adopting coping strategies with flexibility (e.g., Cheng, 2003). Previous research indicates that in the face of certain traumatic events, individuals may rely on reestablishing feelings of personal control/dignity, connectedness, and ▶ solidarity among family and community, as well as reaffirmation of positive, culturally shared values (e.g., religious affirmations) (Pérez-Sales, Cervellón, Vázquez, Vidales, & Gaborit, 2005). Because coping is a dynamic process, being able to employ strategies with flexibility may be one mechanism for achieving resilience in the midst of changing situational demands.

Taken together, the literature indicates that there may not be a "one size fits all" model of resilience. Future research may therefore benefit from exploring and expanding upon different types of resilience pathways and intervention methods. For instance, future research may consider whether factors such as length or quality of resilience training may influence a program's efficacy. Are certain personality types more prone to embrace particular intervention strategies (e.g., positive emotion enhancement) over others? Also, how might gender, culture, or age moderate pathways to resilience and receptivity to varying training methods? Finally, although researchers are currently developing cross-cultural methods of resilience training, further

studies should consider the effects of cultural context on resilience conceptualization (Reivich et al., 2011).

## Cross-References

- ▶ Adaptation
- ▶ Good Neighborhood Index
- ▶ Happiness
- ▶ Health-Related Quality of Life
- ▶ Subjective Well-Being

## References


Arce, E., Simmons, A. N., Stein, M. B., Winkielman, P., Hitchcock, C., & Paulus, M. P. (2009). Association between individual differences in self-reported emotional resilience and the affective perception of neutral faces. *Journal of Affective Disorders, 114*(1–3), 286–293.

Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist, 59*(1), 20–28.

Cheng, C. (2003). Cognitive and motivational processes underlying coping flexibility: A dual-process model. *Journal of Personality and Social Psychology, 84*, 425–438.

DeLongis, A., Coyne, J. C., Dakof, G., Folkman, S., & Lazarus, R. S. (1982). The impact of daily hassles, uplifts and major life events to health status. *Health Psychology, 1*(2), 119–136.

Lyubomirsky, S., & Della Porta, M. (2010). Boosting happiness, buttressing resilience: Results from cognitive and behavioral interventions. In J. W. Reich, A. J. Zautra, & J. Hall (Eds.), *Handbook of adult resilience: Concepts, methods, and applications* (pp. 450–464). New York: Guilford Press.

Masten, A. S. (2007). Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology, 19*(3), 921.

Masten, A. S., & Reed, M. G. J. (2002). Resilience in development. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 74–88). New York: Oxford University Press.

Pérez-Sales, P., Cervellón, P., Vázquez, C., Vidales, D., & Gaborit, M. (2005). Post-traumatic factors and resilience: The role of shelter management and survivors' attitudes after the earthquakes in El Salvador (2001). *Journal of Community & Applied Social Psychology, 15*(5), 368–382.

Reivich, K. J., Seligman, M. E. P., & McBride, S. (2011). Master resilience training in the U.S. Army. *American Psychologist, 66*(1), 25–34.

Skodol, A. E. (2010). The resilient personality. In J. Reich, A. Zautra, & J. Hall (Eds.), *Handbook of adult resilience: Concepts, methods, and applications* (pp. 81–93). New York: Guilford.

Tugade, M. M. (2011). Positive emotions and coping: Examining dual-process models of resilience. In S. Folkman (Ed.), *Oxford handbook of stress health coping* (pp. 186–199). New York: Oxford University Press.

Tugade, M. M., & Fredrickson, B. L. (2004). Resilient individuals use positive emotions to bounce back from negative emotional experiences. *Journal of Personality and Social Psychology, 86*, 320–333.

Tugade, M. M., Fredrickson, B. L., & Barrett, L. F. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality. Special Issue: Emotions, Personality, and Health, 72*(6), 1161–1190.


## Resilient Communities

- ▶ Sustainable Communities Movement
- ▶ Community Adaptation, Arctic

## Resource Management

- ▶ Natural Resource Management (NRM)

## Resource Utilization

- ▶ Life Satisfaction and Sustainable Consumption

R

## Resource-Based Communities


Sasha Tsenkova[1] and Karim Youssef[2]
[1]Faculty of Environment, University of Calgary, Calgary, AB, Canada
[2]University of Calgary, Calgary, AB, Canada


### Synonyms

Resource-dependent communities; Single-industry communities

## Definition

Resource-based communities often depend on a single industry such as forestry, mining, petroleum extraction, or fishing. Closely tied to resource activities, they have unique social and economic characteristics, partly due to their small size, geographical isolation, and integration with the natural environment. Despite community resilience, their quality of life is often challenged by global shifts in resource demand, economic decline, and inability to diversify local economies.

## Description

Resource-based communities (RBCs), also referred to as resource-dependent communities or single-industry communities, are dynamic places where the interface between the natural environment and society is pronounced due to the dependence on a predominant single economic base. Over a thousand RBCs in Canada, home to five million people, continue to be both substantial contributors to the national economy and important components of Canadian identity. Closely tied to resource activities, such as forestry, mining, oil and gas production, and fishing, they exhibit widespread variation in their social fabric and economic specialization (Lucas, 1971). Located usually in peripheral and remote locations, they form part of the core-periphery paradigm of regional systems exposed to multiple external forces: economic downturns, corporate restructuring, shifting global demand for commodities, and changes in federal and provincial policies. While mining and oil and gas extraction dominate northern Quebec and the Prairies, fish and food processing are prominent in Atlantic provinces. Coastal British Columbia and northern Ontario show a more diverse set of industries, which reinforces the need for ▶ local planning responsive to the needs of a distinctive economic base. The economic and social vitality of individual RBCs as well as the quality of life is often at risk due to the "boom" and "bust" cycles in local development affecting employment, traditional lifestyles, and the provision of housing, education/health, retail, and essential infrastructure (water, sewer, and transport).

During the boom phase of development, quality of life is challenged by inadequate infrastructure to accommodate large influxes of labor and the accompanying need for additional housing. Rapid expansion of resource-based industries and related manufacturing activities often leads to environmental degradation with the loss of natural resources, wildlife, agricultural land, forests, and wetlands. On the social side, community life is challenged by shift work, intensive work schedules, and the transient nature of migrant workers. Lack of social services, education, and community facilities to serve a rapidly growing population contributes to further spatial isolation and lower educational attainment (Parkins & Angell, 2011).

The decline phase, often driven by changes in technology, degradation of the resource base, and economic restructuring, profoundly affects quality of life in RBCs. Declining population and high unemployment of male-dominated labor force limits opportunities for adjustment to change and imposes considerable pressure on community livelihoods and well-being (Goldenberg, Shoveller, Koehoorn, & Ostry, 2010). Isolated communities become dependent on seasonal work and government transfers as resources run out, mines or plants become obsolete, and the future remains uncertain. Fluctuations between boom and bust historically have created intense changes in local communities, particularly challenging for long-term development.

Planning to improve quality of life in RBCs centers on three strategic priorities. First, a host of public institutions and programs at the national and provincial level encourage economic diversification in complementary industries and value-added manufacturing (Hessing, Howlett, & Summerville, 2005). Second, tourism and ▶ recreation are often seen as drivers of economic development, particularly in amenity-rich communities with environmental assets (Sander-Regier, McLeman, Brklacich, & Woodrow, 2009). Third, better connectivity on a regional scale allows better access to goods, services, and employment opportunities for

**Resource-Based Communities, Fig. 1** Fort McMurray Resource-Based Community, Regional Municipality of Wood Buffalo, AB (Photography credit: Chris Salvo; Souce: Regional Municipality of Buffalo)

**Resource-Based Communities, Fig. 2** Fort McMurray, AB (Photography credit: Chris Salvo; Source: Regional Municipality of Buffalo)



individual communities, allowing for economies of scale and less disruptive restructuring. Researchers are skeptical regarding the ability of RBCs to diversify their economy, retain their labor force, and survive bust cycles. Enduring "rigidities," due to market imperfections, prevent resource-based economies from adjusting to international market signals resulting in continued economic dependency (Randall & Ironside, 1996). Other coping strategies in different local contexts include public-private partnerships, promoting community ▶ sustainability and financial resource management through community-based groups (Noble & Bronson, 2005). An important element of planning for a more sustainable future of RBCs is the formation of a ▶ place identity (Larsen, 2004). The link between land or sea, traditional resource-based activities, local history, culture, and the natural environment shapes ▶ community values and adds a place-specific dimension to quality of life. This is particularly significant, given the immense economic, social, and environmental diversity of resource-dependent communities in Canada (Figs. 1 and 2).

## References

Goldenberg, S. M., Shoveller, J. A., Koehoorn, M., & Ostry, A. S. (2010). And they call this progress? Consequences for young people living and working in resource-extraction communities. *Critical Public Health, 20*(2), 157–168.

Hessing, M., Howlett, M., & Summerville, T. (2005). *Canadian natural resource and environmental policy: Political economy and public policy* (2nd ed.). Vancouver, BC: UBC Press.

Larsen, S. C. (2004). Place identity in a resource-dependent area of Northern British Columbia. *Annals of the Association of American Geographers, 94*(4), 944–960.

Lucas, R. (1971). *Minetown, Milltown, Railtown: Life in Canadian communities of single enterprise*. Toronto, Canada: University of Toronto Press.

Noble, B. F., & Bronson, J. E. (2005). Integrating human health into environmental impact assessment: Case studies of Canada's northern mining resource sector. *Arctic, 58*(4), 395–405.

Parkins, J. R., & Angell, A. C. (2011). Linking social structure, fragmentation, and substance abuse in a resource-based community. *Community, Work & Family, 14*(1), 39–55.

Randall, J. E., & Ironside, R. G. (1996). Communities on the edge: An economic geography of resource-dependent communities in Canada. *The Canadian Geographer, 40*(1), 17–35.

Sander-Regier, R., McLeman, R., Brklacich, M., & Woodrow, M. (2009). Planning for climate change in Canadian rural and resource-based communities. *Environments, 37*(1), 35–57.

## Resource-Dependent Communities

▶ Resource-Based Communities

## Resourcefulness

▶ Resilience

## Resources Deprivation

▶ Objective and Subjective Deprivation

## Respect for Persons

▶ Human Dignity

## Respect for Persons in Sub-Saharan Cultures

▶ Vitality, Community, and Human Dignity in Africa

## Respect, Generational Concern, and Psychological Well-Being

Sheung-Tak Cheng and Jenny W. H. Chan
Department of Psychological Studies and Center for Psychosocial Health and Aging, Hong Kong Institute of Education, Hong Kong, People's Republic of China

### Definition

Generativity includes generative concern and actions to benefit the next generation. Generative concern is predictive of psychological well-being in later life. Psychological well-being includes positive relations, ▶ purpose in life, ▶ personal growth, and self-acceptance. In modernized societies, the day-to-day interactions between older adults and their children reveal a lack of perceived respect by older adults themselves.

### Description

Erik H. Erikson (1963) delineated eight sequential stages for the human life cycle from infancy to late adulthood. The major psychosocial crisis to be resolved in late adulthood is ego integrity versus despair. Ego integrity is achieved by successful resolution of psychosocial crises of the previous stages. Erikson (1997) further emphasized that generativity is one of the most important factors

to achieve ego integrity. Amid five tasks, namely, generativity, intimacy, identity, autonomy, and ▶ trust, generativity alone accounted for 78 % of the variance in ego integrity (Hannah, Domino, Figueredo, & Hendrickson, 1996). Hence, generativity is crucial for late-life development.

Although generativity is the main developmental task for midlife, it may continue to develop into old age. Erikson (1997) explained that generativity in old age means that older people need to maintain a grand-generative function which is vital for staying alive and the commitment to nurturing younger generations. In contemporary society, generativity may become a major striving in later life. First, many adults still bear their parental responsibility until young old, owing to delay in marriage and ▶ childbearing age. Second, with improved health-care services and increased longevity, older adults can spend more years staying with grandchildren and perform grandparenting roles. Bengtson (2001) stated that multigenerational bonds are more important than those within the nuclear families for the well-being of older adults. Third, older people are encouraged to participate socially. Active engagement with life is a key dimension of successful aging (Rowe & Kahn, 1997), and older adults are encouraged to participate in social and productive activities, such as grandparenting and working as volunteers. Whereas generativity in midlife tends to be dominated by the parenting role, that in later life often goes beyond the family and involve assistance to others in ▶ civic engagement (Keyes & Ryff, 1998). Generativity is an increasingly salient dimension in late-life development.

Different aspects of generativity have been discussed. The 7-facet model developed by McAdams and de St. Aubin (1992) considers *generative concern* as the central concept. Older generations are motivated to benefit the younger generations, in order to fulfill cultural demand and inner desires for immortality. With generative concern, one goes through a process of setting concrete goals and actions, resulting in narrative constructions of the generative self. Younger and midlife adults in Western samples showed that generative concern is moderately to strongly related to psychological well-being. In contrast, generative action is unrelated to well-being even though it is strongly associated with generative concern (Grossbaum & Bates, 2002).

The resource differential between generations decline when the older generation reaches late adulthood (Morgan et al., 1991), and so is the generative capacity (Stewart & Vandewater, 1998). In general, generativity is manifested in social contexts, and its impact on older adults may therefore be partly determined by other's feedback on their generative pursuits. As a result, well-being may be jeopardized if they do not feel being appreciated by others who receive their acts of benevolence. However, it is doubtful whether the feedback from the younger generations is generally positive due to the declining social status of older adults. As a result of rapid social and technological changes and the educational disparity between older and younger generations (the present older cohort went through World War II and had limited educational opportunities), the knowledge older adults possess may be seen as obsolete. As a result, they may not be able to maintain their traditional roles as keepers of traditional wisdom and pass it to the younger generations. McAdams, Hart, and Maruna (1998) named this generation gap as the *generativity mismatch*.

Cheng, Chan, and Chan's (2008) focus group study suggested that the educational disparity between older Chinese adults and younger people prevents older adults from engaging in certain helping roles in the family. On the contrary, older adults are often rejected or criticized by their children on their ways of doing things such as the way they teach their grandchildren. Gradually, they withdraw to passive roles and wait for children's invitation to provide assistance, even in daily household tasks. By doing so, older adults try to maintain their relationship with children in good terms, as minimizing conflicts and preserving harmony are important for sustaining support from children. In fact, older adults are aware of the gap of expertise and ▶ social development between themselves and the next generation. Still, they consider passing on moral and behavioral codes through sharing their hard times in the old days as valuable, and

they want to do so in order to create a more lasting impact. Scholars refer to this as the motive to transcend mortality (Erikson et al., 1986; Erikson, Erikson, & Kivnick, 1984). However, older adults' contributions were often devalued in modernized societies. There is a lack of respect and appreciation in their day-to-day interactions with younger generations. For this reason, Cheng (2009) suggested that attainment of generative goals and well-being are closely related to respect for older adults.

Perceived respect refers to the degree of respect for one's generative acts by the younger generations (Cheng, 2009). ▶ Psychological well-being refers to the eudaimonic view and four subscales, namely, positive relations, purpose in life, personal growth, and self-acceptance, taken from Ryff's (1989) Psychological Well-Being Scale. Cheng's (2009) study tried to explore the relationship between generative acts and well-being with perceived respect from younger generations as a mediator. Because the purpose of generativity is to benefit the younger generations, realizing generative goals would be impossible if they are unwelcomed by younger people including children. In accordance with theories of adaptive aging (Baltes & Baltes, 1990; Brandtstädter & Rothermund, 2002; Heckhausen & Schulz, 1995), perceived disrespect was postulated to result in disengagement from generativity goals (i.e., concern). With generative concern decreasing over time, so would be generative behaviors. Thus, negative feedback from others would result in a downward spiral of the generativity development over time.

Items measuring perceived respect and generative acts were developed from verbatim records in focus group discussions reported in Cheng et al. (2008). The Loyola Generativity Scale (McAdams & de St. Aubin, 1992) was back-translated to provide a measure of generative concern. All measures were internally consistent with expected convergent and discriminant correlations, and the generative act measure was not contaminated by general activity level. Also, perceived respect and generative acts were weakly correlated with social desirability. Items developed for generative acts and perceived respect were specifically concerned with two aspects: civic acts (e.g., volunteer work) and nonspecific acts. Items for nonspecific acts refer to "younger generations" in general but for most older people, those acts would likely have taken place within the family. Data were obtained at two time points, 12 months apart, and construct equivalence was established across time via confirmatory factor analysis. The confirmatory factor analysis also showed that the different measures were tapping different constructs and that the civic and nonspecific domains of acts/respect were sufficiently distinct from each other.

Three main findings are highlighted here. First, the study showed only a moderate level of concern and a relatively low level of action to nurture the younger generations. Furthermore, for the subsample with engagement in civic acts, perceived respect was significantly higher for *civic* acts than for nonspecific ones, and there was also inconsistent evidence suggesting greater involvement in the civic than in the nonspecific domain for this group. Second, the concurrent correlations between generative concern and psychological well-being were moderate. However, different from Western findings (Grossbaum & Bates, 2002), generative act was related to well-being. Generative act was also related to perceived respect which in turn was related to well-being, setting up the conditions for testing the mediating pathway via respect. Third, structural equation model using maximum likelihood estimation showed that at T1, 55–63 % of variance in psychological well-being was explained by the model, whereas it was 70–74 % at T2, suggesting that the model worked very well. The path coefficients showed that respect completely mediated the relationship between generative acts and well-being at baseline. Changes in respect from T1 to T2 also completely mediated the relationship between changes in the frequency of generative acts and changes in well-being. Additionally, T1 perceived respect also predicted T2 generative concern, supporting the hypothesis about goal disengagement. Reductions in T2 generative concern as a result of T1 respect led further to decreases in generative acts and well-being over time, supporting the downward spiral hypothesis.

## Discussion

Cheng's (2009) study investigated the relationship between generativity and well-being and the mediating role of respect from younger generations. He asked the following question: "Why should one continue to attempt to guide the younger generations when such attempts are not welcome?" The study found two pathways by which perceived respect affect generativity development and well-being. First, it mediated completely the relationship between generative acts and psychological well-being. In other words, generative acts result in enhanced well-being only because it tends to lead to respect. Those who are able to engage in more generative acts may have more resources and what Stewart and Vandewater (1998) called the generative capacity. These individuals are likely to be respected, especially in the civic domain, and it is the feeling of respect that is crucial in the promotion of psychological well-being. The second pathway by which perceived respect influences generativity development and well-being relates to its effect on the disengagement from generative goals, leading to a downward spiral in generative acts and well-being over time. Cheng (2009) found that even after controlling for T1 generative concern, T2 generative concern was significantly predicted by T1 perceived respect, though the effect was a relatively small one, suggesting that the disengagement from generative goals due to a lack of respect was probably a gradual process.

Older adults felt more respect for civic acts than for nonspecific acts (mostly in the family domain in reality). Their helping behaviors in the family may be taken for granted, and hence an open display of respect may not have taken place frequently enough. In fact, their helping behaviors in the family may often be criticized or rejected by their offspring (Cheng et al., 2008). On the contrary, generative behaviors in the civic domain, such as voluntary work, are often recognized formally one way or another. There may be another dimension to the interactions between older parents and their adult children in this regard. Showing respect to older parents is a core aspect of filial piety in Chinese/Asian cultures. It is also an important factor determining older people's well-being as well as whether they feel their children have honored their filial obligations (Cheng & Chan, 2006; Cheng & Chan, in press). The lack of respect for older people's generative behaviors in the family may contribute further to declines in well-being through disappointments with children's filial piety. This needs to be addressed in future research.

Lastly, it should be mentioned that generativity is carried out in particular social and cultural environments that include appropriate normative expectations and chances for guiding younger people (McAdams & de St. Aubin, 1992). Whether the role of respect is specific to Chinese and Asian cultures which emphasize respect for older people is not clear. Cross-cultural research is needed to determine the extent to which Cheng's (2009) findings are generalizable to other cultural contexts.

## References


Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). Cambridge, MA: Cambridge University Press.

Bengtson, V. L. (2001). Beyond the nuclear family: The increasing importance of multigenerational bonds. *Journal of Marriage and Family, 63*, 1–16.

Brandtstädter, J., & Rothermund, K. (2002). The life-course dynamics of goal pursuit and goal adjustment: A two-process framework. *Developmental Review, 22*, 117–150.

Cheng, S. T. (2009). Generativity in later life: Perceived respect from younger generations as a determinant of goal disengagement and psychological well-being. *Journal of Gerontology: Psychological Sciences, 64B*(1), 45–54.

Cheng, S.-T., & Chan, A. C. M. (2006). Filial piety and psychological well-being in well older Chinese. *Journal of Gerontology: Psychological Sciences, 61B*, P262–P269.

Cheng, S.-T., Chan, W., & Chan, A. C. M. (2008). Older people's realisation of generativity in a changing society: The case of Hong Kong. *Ageing & Society, 28*, 609–627.

Cheng, S. -T., & Chan, J. W. H. (in press). Wellbeing and filial piety. In A. C. Michalos (Ed.), *The encyclopedia of quality of life research*. Berlin, Germany: Springer.

Erikson, E. (1963). *Childhood and society* (2nd ed.). New York: Norton. (Original work published 1950)
Erikson, E. (1997). *The life cycle completed: Extended version with new chapters on the ninth stage of development by Joan M. Erikson*. New York: Norton.
Erikson, E. H., Erikson, J. M., & Kivnick, H. Q. (1986). *Vital involvement in old age*. New York: Norton.
Grossbaum, M. F., & Bates, G. W. (2002). Correlates of psychological well-being at midlife: The role of generativity, agency and communion, and narrative themes. *International Journal of Behavioral Development, 26*, 120–127.
Hannah, M. T., Domino, G., Figueredo, A. J., & Hendrickson, R. (1996). The prediction of ego integrity in older persons. *Educational and Psychological Measurement, 56*, 930–950.
Heckhausen, J., & Schulz, R. (1995). A life-span theory of control. *Psychological Review, 102*, 284–304.
Keyes, C. L. M., & Ryff, C. D. (1998). Generativity in adult lives: Social structural contours and quality-of-life consequences. In D. P. McAdams & E. de St. Aubin (Eds.), *Generativity and adult development*. Washington, DC: American Psychological Association.
Kotre, J. (1984). *Outliving the self: How we live on in future generations*. Baltimore, MD: Johns Hopkins University Press.
McAdams, D. P., & de St. Aubin, E. (1992). A theory of generativity and its assessment through self-report, behavioral acts, and narrative themes in autobiography. *Journal of Personality and Social Psychology, 62*, 1003–1015.
McAdams, D. P., Hart, H. M., & Maruna, S. (1998). The anatomy of generativity. In D. P. McAdams & E. de St Aubin (Eds.), *Generativity and adult development: How and why we care for the next generation* (pp. 7–43). Washington, DC: American Psychological Association.
Morgan, D. L., Schuster, T. L., & Butler, E. W. (1991). Role reversals in the exchange of social support. *Journal of Gerontology, 46*, S278–S287.
Rowe, J. W., & Kahn, R. L. (1997). Successful aging. *The Gerontologist, 37*, 433–440.
Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*, 1069–1081.
Stewart, A. J., & Vandewater, E. A. (1998). The course of generativity. In D. P. McAdams & E. de St. Aubin (Eds.), *Generativity and adult development: How and why we care for the next generation*. Washington, DC: American Psychological Association.

## Respondent Abuse

► Respondent Burden

## Respondent Burden

Christopher McCarty
UF Survey Research Center, University of Florida Bureau of Economic and Business Research, Gainesville, FL, USA

### Synonyms

Respondent abuse

### Definition

The impact of the interviewing process on individual respondents.

### Description

Concerns about interview length and frequency date back to the 1920s. With the steady increase in the number and size of federally sponsored surveys, the US Office of Management and Budget set policies to reduce respondent burden with the Commission on Federal Paperwork (Sharp & Frankel, 1983). Many of these policies are concerned with eliminating redundant surveys. Guidelines for the protection of human subjects and institutional review boards are closely linked with the reduction of respondent burden. Respondent characteristics, such as age and level of education, may make the same survey burdensome to one group and not another. Wenemark, Frisman, Svensson and Kristenson (2010) showed a relationship between perceptions of burden and a measure of respondent motivation from intrinsic to extrinsic.

Bradburn (1979) proposed four categories of respondent burden – interview length, respondent effort required, interview frequency, and stress from interview content. Interview length is associated with survey break-off (Peytchev, 2009) which is more common in Web surveys than telephone or face-to-face interviews. Interviews of more than 30 min are typically considered by researchers to

be burdensome. Interview frequency is of particular concern for small populations that receive a disproportionate amount of public assistance from programs that require evaluations. Public agencies that fund such surveys can reduce the burden by coordinating data collection. Panel surveys where respondents are interviewed over time may result in attrition associated with the burden of repeated interviews. However, Deeg, van Tilburg, Smit and de Leeuw (2002) showed that respondents exposed to repeated side studies split into one set who dropped out early and the remainder who had a higher tolerance for burden. Some respondents may consider survey content to be a burden, such as surveys on sexual attitudes and behavior. This burden may be reduced by using modes that allow respondents to enter answers themselves, such as paper and pencil, Web, or computer-assisted personal interviewing (CAPI). Questionnaire design can also make interviews more or less burdensome (Jäckle, 2008). The skill level of the interviewer such as ability to speak clearly and maintaining a positive attitude may compensate for factors that would otherwise be burdensome. Respondent incentives, typically in the form of cash or gift cards, are common, particularly with cell phone interviews where respondents pay for minutes used. Survey research is increasingly conducted through Web interfaces, on cell phones, and through social media. The expectations of respondents regarding what constitutes appropriate burden with these media will require a different perspective (Peytchev, Couper, Mccabe, & Crawford, 2006). Emerging topics, such as social network analysis, will involve new types of burden (McCarty & Killworth, 2007).

## Cross-References

- ▶ Burnout
- ▶ Data Collection Methods
- ▶ Fatigue
- ▶ Follow-Up Questions
- ▶ Mode Effects
- ▶ Probing Questions or Probes
- ▶ Questionnaire Design
- ▶ Response Bias(es)
- ▶ Response Rate
- ▶ Sample Survey
- ▶ Social Indicators
- ▶ Survey Research
- ▶ Web-Based Questionnaire

## References

Bradburn, N. (1979). Respondent burden. In L. Reeder (Ed.), *Health survey research methods: Second biennial conference* (Williamsburg, VA). Washington, DC: U.S. Government Printing Office.

Deeg, D. J. H., van Tilburg, T., Smit, J. H., & de Leeuw, E. D. (2002). Attrition in the longitudinal aging study Amsterdam: The effect of differential inclusion in side studies. *Journal of Clinical Epidemiology, 55*, 319–328.

Jäckle, A. (2008). Dependent interviewing: Effects on respondent burden and efficiency of data collection. *Journal of Official Statistics, 24*(3), 411–430.

McCarty, C., & Killworth, P. D. (2007). Impact of methods for reducing respondent burden on personal network structural measures. *Social Networks, 29*, 300–315.

Peytchev, A., Couper, M. P., Mccabe, S. E., & Crawford, S. D. (2006). Web survey design: Paging versus scrolling. *Public Opinion Quarterly, 70*(4), 596–607.

Peytchev, A. (2009). Survey breakoff. *Public Opinion Quarterly, 73*(1), 74–97.

Sharp, L. M., & Frankel, J. (1983). Respondent burden: A test of some common assumptions. *Public Opinion Quarterly, 47*(1), 36–53.

Wenemark, M., Frisman, G. H., Svensson, T., & Kristenson, M. (2010). Respondent satisfaction and respondent burden among differently motivated participants in a health-related survey. *Field Methods, 22*(4), 378–390.



# Response Bias(es)

Peter B. Smith
School of Psychology, University of Sussex, East Sussex, UK

## Definition

Response biases occur when respondents complete rating scales in ways that do not accurately reflect their true responses. They occur especially among responses to Likert

scales that ask the respondent to agree or disagree with various statements.

## Description

Ratings may be affected by the respondent's wish to provide a response that is socially desirable. In differing contexts this may cause them to make ratings toward the positive end of rating scales (acquiescence), toward the extremes of scales (extremity) or toward the middle of scales (moderation). Responses may also be biased by the sequencing of questions in a survey. Response bias can be detected and controlled in any of three ways. Firstly, half of the items referring to a topic in a survey can be worded in the reverse direction. Secondly, a locally valid measure of social desirability can be included in the survey. Thirdly, additional unrelated items can be included in a survey that focus on topics where it is not logically possible for a respondent to agree (or to disagree) with all the items. Estimates of response bias can then be partialled out of mean survey responses. Where only a single survey item is used to measure an aspect of quality of life, it is impossible to determine the presence or absence of response bias, and such a measure is unlikely to be valid.

Response biases are particularly problematic when comparisons are made between mean scores obtained from samples drawn from different cultural or national groups (Johnson, Shavitt, & Holbrook, 2011). Cultural groups within which there are norms favoring harmony show greater acquiescence and greater moderation than is found in cultural groups where expression of individual distinctiveness is normative (Smith, 2004, 2011; Suh, Diener, Oishi, & Triandis, 1998). In more collectivistic populations, extremity bias is greater when there are a larger number of response options (Hui & Triandis, 1989).

Persons in individualistic cultures (such as those in northern Europe and North America) predominantly favor analytic cognition, whereas persons in more collectivistic cultures (such as those in East Asia) more often employ holistic cognition (Nisbett, Peng, Choi, & Norenzayan, 2000). This means that respondents in individualistic cultures will react to each item in a survey without being strongly affected by other items in a survey. In contrast, respondents in collectivistic cultures will be more attentive to context, and their responses to a given item will be affected by the nature of the preceding items. For instance, Haberstroh, Oyserman, Schwarz, Kuhnen, and Ji (2001) compared the effects of question sequence among German and Chinese students. When a question on life satisfaction preceded a question on academic satisfaction, the two responses were moderately correlated in both samples. When the sequence was reversed, the correlation between the two responses was markedly reduced among the Chinese. Haberstroh et al. inferred that after responding to the question about academic satisfaction, the Chinese respondents would have assumed that the question about life satisfaction referred to other aspects of satisfaction.

## References

Haberstroh, S., Oyserman, D., Schwarz, D., Kuhnen, U., & Ji, L. (2001). Is the interdependent self more sensitive to question context than the independent self? *Journal of Experimental Social Psychology, 38*, 323–329.

Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology, 20*, 296–309.

Johnson, T. P., Shavitt, S., & Holbrook, A. (2011). Survey response styles across cultures. In D. Matsumoto & F. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 130–175). New York: Cambridge University Press.

Nisbett, R., Peng, K. P., Choi, I., & Norenzayan, A. (2000). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*, 291–310.

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology, 35*, 50–61.

Smith, P. B. (2011). Communication styles as dimensions of national culture. *Journal of Cross-Cultural Psychology, 42*, 216–233. doi:10.1177/0022022110396866.

Suh, E. M., Diener, E., Oishi, S., & Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology, 74*, 482–493.

# Response Format

Eric K. H. Chan
Department of ECPS, University of British Columbia, Vancouver, BC, Canada

## Synonyms

Interval response format; Item analysis; Item response theory; Likert scale; Nominal response; Psychometric methods; Ratio response format; Scaling

## Definition

The three most commonly used response formats in quality of life measures include (1) binary, (2) ordinal, and (3) continuous. With the binary response format, only two possible values are involved (e.g., Yes or No). The ordinal response format typically consists of three to ten response values (e.g., 5-point Likert scale). In the continuous response format, an infinite number (or very large number) of score points can be produced.

## Description

Researchers developing a quality of life measure need to choose a response format for the measure. The most widely used response formats in quality of life and related measures include (1) binary (dichotomous), (2) ordinal, and (3) continuous.

With the binary response format, the quality of life variables of interest consist of only two possible values, such as yes or no, like or dislike, and happy or not happy. For example, respondents are asked to answer true or false to a series of items measuring happiness. The multiple choice format, although rarely seen in quality of life measures, is also a form of binary response format as it is scored as either correct or incorrect. To learn more about the binary response format, please refer to the Binary Response Format entry in this volume.

In the social and behavioral sciences, the ordinal response format is perhaps the most commonly used. Measures using ordinal responses typically consist of three to ten response values. For example, in a 5-point Likert scale, respondents are provided with options such as "*Completely Untrue*, *Mostly Untrue*, *Equally Untrue and True*, *Mostly True*, *Completely True*" or "*Strongly Disagree*, *Disagree*, *Neither Disagree* or *Agree*, *Agree*, *Strongly Disagree*." The literature is replete with empirical studies of issues such as the influence of the number of response options on item responses, the advantages and disadvantages of even and odd number of points, the distance between points, and the cognitive burden associated with having a "neutral" point (e.g., "Neither Agree nor Disagree"). Measurement specialists have yet to come to a consensus on the best practices for the ordinal response format.

The continuous response format produces scores with an infinite (or very large number) of points. Although some researchers view items with a large number of response points (e.g., 7-point Likert scale) as practically continuous, perhaps the type of response format that is arguably the closest to a true continuous scale is the visual analogue scale (VAS). In VAS, respondents are presented with a continuous horizontal or vertical line with, for example, two anchored endpoints or descriptors representing the opposite ends of the continuum. Respondents are required to place a mark on the line that represents their response to an item. Prior to the ubiquitous availability of computers, the scoring of VAS was a tedious task (it required measuring the distance using a ruler for each item, for each respondent). With the advancement in computer technology, we may see more quality of life measures employing VAS.

Compared to ordinal and continuous response formats, employing a binary response format places less cognitive burden on respondents and may be particularly useful when there are many items in the scale. However, the binary response format tends to lose information forcing respondents to endorse one of the two options. Different response formats also require the use of different statistical and measurement methods. In general, larger sample sizes are required to run psychometric analyses on ordinal scales than binary

scales. In the examination of internal consistency, different response formats require different psychometric procedures (Gadermann, Guhn, & Zumbo, 2012; Zumbo, Gadermann, & Zeisser, 2007). Response formats also have an influence on differential item functioning (DIF) analysis (Gelin & Zumbo, 2003). In addition, a series of studies have demonstrated that outliers across different response formats have an influence on the internal consistency of measures (Liu & Zumbo, 2007; Liu, Wu, & Zumbo, 2010).

### Cross-References

- ▶ Item Analysis
- ▶ Item Response Theory
- ▶ Likert Scale
- ▶ Ordinal Alpha

### References

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research and Evaluation, 17*(3), 1–13.

Gelin, M. N., & Zumbo, B. D. (2003). DIF results may change depending on how an item is scored: An illustration with the center for epidemiological studies depression (CES-D) scale. *Educational and Psychological Measurement, 63*, 65–74.

Liu, Y., Wu, A. D., & Zumbo, B. D. (2010). The impact of outliers on Cronbach's coefficient alpha estimate of reliability: Ordinal/rating scale item responses. *Educational and Psychological Measurement, 70*, 5–21.

Liu, Y., & Zumbo, B. D. (2007). The impact of outliers on Cronbach's coefficient alpha estimate of reliability: Visual analogue scales. *Educational and Psychological Measurement, 67*, 620–634.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods, 6*, 21–29.

## Response Options Bias

- ▶ Life Satisfaction Ratings and Response Formats

## Response Pattern Approach to Identifying IER

- ▶ Survey Responses with Insufficient Effort

## Response Rate

- ▶ Survey Research

## Response Shift

Carolyn E. Schwartz[1] and Mirjam A. G. Sprangers[2]
[1]DeltaQuest Foundation, Tufts University Medical School, Concord, MA, USA
[2]Department of Medical Psychology, University of Amsterdam, Amsterdam, The Netherlands

### Synonyms

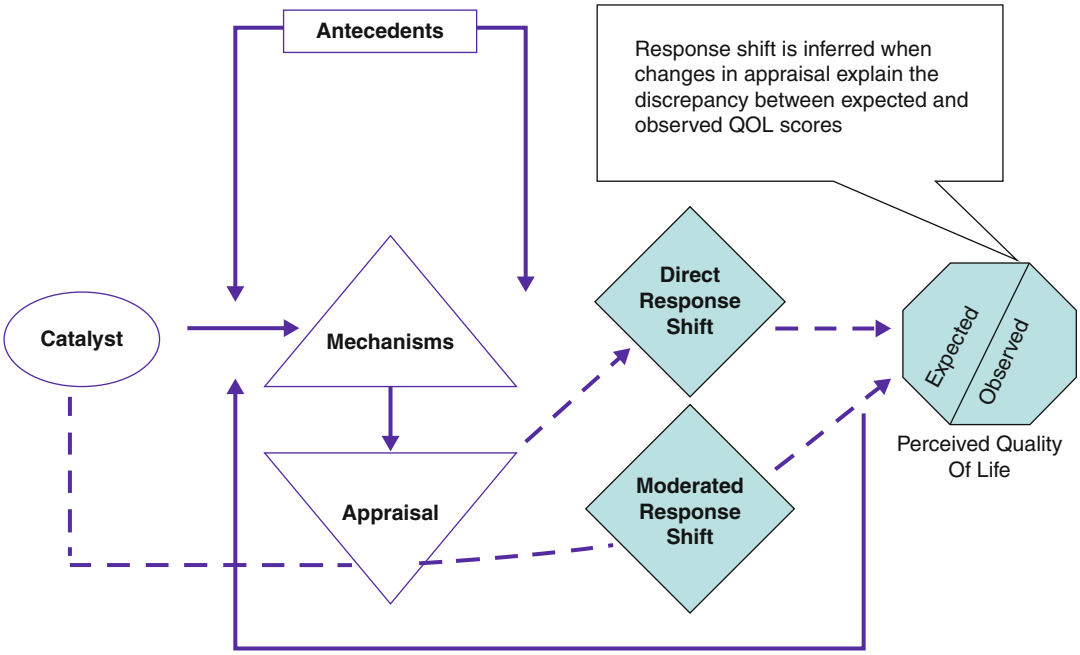
Adaptation; Appraisals and response shifts; Recalibration response shift; Reconceptualization response shift; Reprioritization response shift

### Definition

Response shift refers to a change in the meaning of one's self-evaluation of a target construct as a result of (a) a change in the respondent's internal standards of measurement (scale recalibration), (b) a change in the respondent's values (reprioritization), or (c) a redefinition of the target construct (reconceptualization).

### Description

The term "response shift" was coined by Howard and colleagues and originates from research on educational training interventions (Howard et al., 1979). They defined response shift in terms




Adapted from Sprangers and Schwatz (1999) and Rapkin and Schwartz (2004) models of appraisal and quality of life (permission pending).

**Response Shift, Fig. 1**

of changes in internal standards of measurement. Parallel in time but independently, Golembiewski et al. forwarded a change typology originating from management sciences (Golembiewski, Billingsley, & Yeager, 1976). They distinguished observable change (alpha change) from changes in internal standards (beta change) and meaning (gamma change) (Golembiewski et al., 1976). Changes in values were inherent to Golembiewski's description of reconceptualization, but were not explicitly seen as a separate component. These concepts lead to the above *working definition* proposed by Sprangers and Schwartz (2000). Reprioritization is thus an added explicit component of this definition.

This working definition was integrated into a theoretical model that relates health changes to QOL. In this model, changes in health status (*catalysts*) may induce response shift via *mechanisms* to accommodate this change. Mechanisms refer to cognitive, affective, and behavioral strategies to cope, such as engaging in a special type of social comparison or seeking social support. The kinds of mechanisms individuals may use depend not only on the specific health change but also on stable characteristics of the person, the so-called antecedents, such as gender and personality traits. Response shift affects one's self-evaluation of QOL (Fig. 1). This model launched a host of theory-driven studies, but further development of the model was required to rectify the mechanism-response-shift circularity: both terms were operationalized similarly.

Efforts to hone the QOL response-shift construct and model have focused on further delineating the three response-shift aspects in terms that were complementary and provided possible solutions to the above-noted circularity. A first refinement was developed by Rapkin and Schwartz (2004) by conceptualizing response shift as an "epiphenomenon"; rather than being measured directly, it is inferred when changes in *appraisal* explain discrepant (e.g., residual) changes in QOL scores that cannot be explained by the "standard QOL model" (Fig. 1).

This "standard QOL model" has as a primary hypothesis that catalysts – changes in health, treatment, and life events – are directly related to QOL and that the effects of antecedents (demographic factors, personality, cultural and historical influences) on QOL are mediated through catalysts (Rapkin & Schwartz, 2004). Their model adopts four appraisal parameters that map to the three types of response shift [1]: changes in the *frame of reference* relate to *reconceptualization*, changes in *strategies for sampling experience* within one's frame of reference deemed relevant to rating QOL as well as changes in the factors that determine the relative salience of different experiences relate to *reprioritization*, and changes in *standards of comparison* for evaluating one's experience relate to *recalibration*. The observed QOL score and change in this score over time is determined by an internal subjective process, a *combinatory algorithm*, which reflects how individuals understand and appraise QOL. The *contingent true score* can be validly compared across or within individuals and over time only if the appraisal parameters are similar (Schwartz & Rapkin, 2004).

A second refinement put forward by Oort and colleagues built on existing distinctions and statistical definitions of measurement bias to distinguish response shift from two complementary perspectives: measurement and conceptual perspectives (Oort, 2005; Oort, Visser, & Sprangers, 2009). Both ways of framing response shift formulate it in statistical terms as a *violation of conditional independence*. In the measurement perspective, response shift is conceptualized as measurement invariance over time. That is, changes in *observed QOL scores* cannot be fully explained by true changes in QOL. In the conceptual perspective, response shift is conceptualized as the effect of other cognitive variables on change in *true/expected QOL* that are different from the acknowledged explanatory variables. In both the measurement and conceptual perspectives, bias is created by not considering sources of variance that affect measurement or explanation of QOL changes, respectively (Oort, 2005; Oort et al., 2009).

Both refinements make a further important distinction: that response shift is not a variable but rather an *effect*. Rapkin and Schwartz (2004) consider response shift to be something that is inferred in a specific context (i.e., changes in appraisal explain residual variance in QOL over time). The Oort approach also adopts the idea that response shift is inferred when specific effects are detected, either in the measurement or the explanation of QOL change over time (Oort et al., 2009). Characterizing response shift as an effect rather than as a construct one measures directly has implications for how researchers would operationalize it and report their findings, i.e., response-shift effects, not response shifts.

Response shift has now been studied and recognized to affect adaptation to a wide degree of health conditions, including multiple sclerosis, cancer, stroke, diabetes, geriatrics, palliative care, dental disorders, and orthopedics. A meta-analysis done on response shift reported that response shift findings ranged from moderate to small effect sizes (Schwartz et al., 2006); thus, adjusting for response shift in the data analytic phase of a clinical study can boost effect sizes from moderate to large (Oort, Visser, & Sprangers, 2005). Other research has demonstrated that adjusting for response shift can even reverse putative null or deleterious findings (Schwartz, Feinberg, Jilinskaia, & Applegate, 1999).

## Response-Shift Methods

Current methods for detecting response shifts are evolving from a predominant focus on the *retrospective pretest-posttest design (then-test) approach* (Howard et al., 1979; Sprangers et al., 1999) to place an emphasis on statistical or individualized methods. The retrospective pretest-posttest design aims to detect change in internal standards or "recalibration response shift" by extending the standard pretest-posttest design by readministering questionnaire(s) at the time of follow-up, most times immediately after completion of the posttest. To that purpose, the instruction is adapted such that respondents are asked to think back to how they were doing at the start of the study (e.g., pretreatment) and to

answer a given questionnaire retrospectively, thereby reevaluating their QOL of that time. Then-test-minus-pretest difference scores would represent a recalibration response-shift effect, and posttest-minus-pretest would represent the adjusted treatment or time effect. The approach assumes that the posttest and then-test share the same internal standards and their mean change would thus allow a better estimate of a treatment or time effect than a standard posttest-minus-pretest difference score. The "then-test" approach has the advantage of being easy to administer and analyze but the disadvantages of random error and/or confounding with recall bias as well as being difficult to interpret. For these reasons, we now briefly describe promising statistical or individualized methods that have evolved in the first decade of the twenty-first century.

There are three statistical methods and one mixed method (i.e., qualitative and quantitative) that have been applied to response-shift detection that hold promise: structural equation modeling, latent trajectory analysis with subject-centered residuals, classification and regression tree analysis, and the SEIQOL individualized method. The three statistical methods require substantial sample sizes, on the order of ten subjects per variable and a minimum of 200 subjects. The four methods vary in terms of how much they focus on aggregate analyses versus individual patient-focused analysis and thus how sensitive they are to individual response shifts.

Originally evolving from factor analytic methods, structural equation modeling is a technique that combines factor analysis and regression analysis to solve multivariate research questions at a group level (Oort et al., 2005). By analyzing covariance matrices, these models test measurement and structural models to first test the assumption of measurement invariance and then to examine whether relationships among variables are similar over time (i.e., the structural model). Recent advances of this method were made by Oort et al. (2005), to clarify how distinct changes detectable with structural equation modeling reflect different aspects of response shift. This work extended earlier work done by Schmitt (1982) and yielded more sensitive algorithms for detecting response shifts. The Oort method has been applied to cancer, stroke, musculoskeletal diseases, and chronic obstructive pulmonary disease patients.

Latent trajectory analysis with subject-centered residuals is a method developed by Mayo et al. that focuses on the individual and seeks to develop a predictive growth curve model to examine patterns in discrepancies between expected and observed scores (Mayo, Scott, & Ahmed, 2009). By obtaining and scaling model residuals, Mayo creates subject-centered residuals to categorize respondents as either (1) exhibiting no response shift, i.e., the person's residuals are consistent over time, but there was some change in their perceived QOL at time t; (2) exhibiting a positive response shift, i.e., the person's evaluation started low and then shifted or reassessed upward; and (3) exhibiting a negative response shift, i.e., started higher than expected and then reassessed down over time. This method has been applied to stroke and multiple sclerosis patients.

Classification and regression tree analysis (Breiman, Friedman, Olshen, & Stone, 1993) is a method applied by Li and Rapkin (2009) that combines qualitative and quantitative methods to yield a rich analysis of complex data. These investigators utilized the QOL Appraisal Profile developed by Rapkin and Schwartz (2004), which yields qualitative text data in response to open-ended questions as well as quantitative data in response to multiple choice questions. The tool measures four distinct parameters of the appraisal process: (a) *Framing* (i.e., what does QOL mean to the individual?), (b) *Sampling* (i.e., what relevant experiences do I have?), (c) *Evaluating* (i.e., how do experiences compare to relevant standards?), and (d) *Combining* (i.e., what is the relative importance of different experiences?). These data are then content analyzed to yield categories amenable to quantitative analysis, and tree-like diagrams are generated that show the variable cut-off "paths" for subgroups of respondents. The final product of this analysis is homogenous groupings of respondents who share patterns of appraisal.

Individualized methods are another novel approach to evaluating response shifts. One of

the most familiar is the ▶ Schedule for the Evaluation of Individual QOL (SEIQOL) (Joyce, 1999; Ring, Hofer, Heuston, Harris, & O'Boyle, 2005), which has been used to examine stability in domains and internal metrics (i.e., cues and anchors) over time and across groups. The three aspects of response shift (i.e., recalibration, reconceptualization, and reprioritization) are thus explicitly included in the SEIQOL method, making it an appealing and interpretable approach to getting highly individualized metrics of response shift. This method involves a semi-structured interview that elicits from the person the domains or areas she/he thinks about when considering QOL ("cues"), the response options that make the most sense to the person ("anchors"), and his/her level of functioning within each of these domains ("levels"). The person's overall score is generated using a regression model approach. At each subsequent interview, domains, cues, and levels are elicited from the person and an overall score is generated. The result is an idiographic (i.e., individual-focused) QOL score that allows the person's domains, cues, and levels to vary as they do naturally. This individualized QOL tool has been found to be useful in understanding changes in individuals over time as they cope with a progressive and deteriorating disease and could be useful to elucidate how this process is similar or different among people with a substantial albeit stable disabling condition.

## Cross-References

- ▶ Changes in Quality of Life
- ▶ SEIQoL: Schedule for the Evaluation of Individual Quality of Life

## References

### Literature Cited

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and regression trees*. New York: Chapman & Hall/CRC.

Golembiewski, R. T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal Applied Behavior Science, 12*, 133–157.

Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychology Measurement, 3*(1), 1–23.

Joyce, C. R. B., O'Boyle, C. A., & McGee, H. (1999). *Individual quality-of-life: Approaches to conceptualisation and assessment*. Amsterdam: Harwood Academic.

Li, Y., & Rapkin, B. (2009). Classification and regression tree analysis to identify complex cognitive paths underlying quality-of-life response shifts: A study of individuals living with HIV/AIDS. *Journal of Clinical Epidemiology, 62*, 1138–1147.

Mayo, N., Scott, C., & Ahmed, S. (2009). Case management post-stroke did not induce response shift: The value of residuals. *Journal of Clinical Epidemiology, 62*, 1148–1156.

Oort, F. J. (2005). Towards a formal definition of response shift. *Quality of Life Research, 14*, 2353–2355.

Oort, F. J., Visser, M. R. M., & Sprangers, M. A. G. (2005). An application of structural equation modeling to detect response shifts and true change in quality-of-life data from cancer patients undergoing invasive surgery. *Quality of Life Research, 14*, 599–609.

Oort, F. J., Visser, M. R. M., & Sprangers, M. A. G. (2009). Measurement and conceptual perspectives on response shift: Formal definitions of measurement bias, explanation bias, and response shift. *Journal of Clinical Epidemiology, 62*, 1126–1137.

Rapkin, B. D., & Schwartz, C. E. (2004). Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health and Quality of Life Outcomes, 2*(1), 14.

Ring, L., Hofer, S., Heuston, F., Harris, D., & O'Boyle, C. A. (2005). Response shift masks the treatment impact on patient reported outcomes (PROs): The example of individual quality-of-life in edentulous patients. *Health and Quality of Life Outcomes, 3*, 55.

Schmitt, N. (1982). The use of analysis of covariance structures to assess beta and gamma change. *Multivariate Behavioral Research, 17*, 343–358.

Schwartz, C. E., & Rapkin, B. D. (2004). Reconsidering the psychometrics of quality-of-life assessment in light of response shift and appraisal. *Health and Quality of Life Outcomes, 2*, 16.

Schwartz, C. E., Feinberg, R. G., Jilinskaia, E., & Applegate, J. C. (1999). An evaluation of a psychosocial intervention for survivors of childhood cancer: Paradoxical effects of response shift over time. *Psycho-Oncology, 8*(4), 344–354.

Schwartz, C. E., & Sprangers, M. A. G. (2000). *Adaptation to changing health: Response shift in quality-of-life research*. Washington, DC: American Psychological Association.

Schwartz, C., Bode, R., Repucci, N., Becker, J., Sprangers, M., & Fayers, P. (2006). The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research, 15*, 1533–1550.

Sprangers, M. A., Van Dam, F. S., Broersen, J., Lodder, L., Wever, L., Visser, M. R., et al. (1999). Revealing response shift in longitudinal research on fatigue – The use of the thentest approach. *Acta Oncologica, 38*(6), 709–718.

## Response Style Effects

▶ Life Satisfaction Ratings and Response Formats

## Response Time Approach to Identifying IER

▶ Survey Responses with Insufficient Effort

## Response to Stress

▶ Stress Reactivity

## Response-Time Modeling

▶ Event History Analysis

## Responsibility

▶ Personal and Institutional Accountability

## Responsible Conduct of Research

▶ Research Integrity

## Responsible Treatment of Participants in Research

▶ Research Ethics

## Responsiveness to Change

Caroline B. Terwee
Department of Epidemiology and Biostatistics, VU University Medical Center, EMGO Institute for Health and Care Research, Amsterdam, The Netherlands

### Synonyms

Longitudinal validity; Sensitivity; Sensitivity to change

### Definition

We adopt the definition of responsiveness of the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) group, who defined responsiveness as "the ability of an instrument to detect change over time in the construct to be measured" (Mokkink et al., 2010). Responsiveness is considered as an aspect of validity.

### Description

According to the COSMIN taxonomy, validity refers to the validity of a single score, while responsiveness refers to the validity of a change score. Responsiveness is only relevant for measurement instruments that are used in evaluative applications, that is, when the instrument is used in a longitudinal study to measure change over time.

When assessing responsiveness, the hypothesis is tested that if patients change on the construct of interest, their score on the measurement

instrument assessing this construct should change accordingly. Therefore, in analogy to validity, assessing responsiveness consists of testing hypotheses about the expected relationships between changes on the instrument under study and changes on other instruments that measure similar or different constructs or expected differences between groups in changes on the instrument. For assessing responsiveness, a longitudinal study with repeated measurements is required. The time period should be chosen in such a way that it can be expected that during this time period, at least a proportion of the patients will improve or deteriorate on the construct to be measured. The study population should be representative for the target population and situation in which the instrument will be used in future studies or clinical practice. One should be cautious when evaluating the responsiveness of a measurement instrument in the same study in which the instrument is used as an outcome measure because results on the quality of the instrument and the effect of the intervention cannot be disentangled (de Vet, Terwee, Mokkink, & Knol, 2011).

Assessing responsiveness is a continuous process of accumulating evidence. It is not possible to formulate standards for the number of hypotheses that need to be tested. This depends on the construct to be measured, the study population and context, and the content and measurement properties of the instruments used for comparison. One can therefore never conclude that an instrument *is* responsive. Different approaches can be used, and the evidence from these approaches should be combined in order to draw conclusions about the degree of responsiveness of the instrument in a specific population and context. The two main approaches for assessing responsiveness are the criterion approach and the construct approach.

## Criterion Approach

When a gold standard for change is available, a criterion approach can be used to assess the degree to which changes in the scores on a measurement instrument are an adequate reflection of changes in scores on a "gold standard." This is comparable to assessing criterion validity. In theory, the gold standard is a perfectly valid assessment. It is usually a measurement instrument that has been accepted as a gold standard by experts. For the measurement of quality of life and other Patient-Reported Outcomes (PROs), a gold standard does not exist. An exception is a situation in which a short questionnaire is compared to a long version. In that case, one might consider the long version as the gold standard.

In many studies on the responsiveness of quality of life instruments, a global rating scale (GRS) is used as a gold standard for measuring change. Patients are asked at follow-up, in a single question, to indicate how much they have changed, for example, on a 5-point rating scale ranging from much worse to much better. Such a GRS has a high face validity, but doubt has also been expressed about the reliability and validity of such retrospective measures of change (Norman, Stratford, & Regehr, 1997). However, if the GRS measures the same construct as the instrument under study, it may be considered a reasonable gold standard for quality of life instruments.

Agreement between changes in scores on the measurement instrument under study and changes in scores on the gold standard can be assessed by calculating correlation coefficients. Other commonly used methods are sensitivity and specificity, limits of agreement, and Intraclass Correlation Coefficients (ICCs). The area under the Receiver Operating Curves (ROC) (AUC) is also often used, which refers to the ability of an instrument to discriminate between patients who are considered to be improved and patients who are not considered to be improved according to the gold standard.

It is best to decide a priori which level of agreement between changes in the scores on the measurement instrument and changes in scores on the gold standard one considers acceptable. Correlations above 0.7 are sometimes reported to be acceptable. An AUC of at least 0.70 is usually considered to be appropriate.

## Construct Approach

If there is no gold standard available, the assessment of responsiveness relies on testing

hypotheses, just like the assessment of construct validity. The hypotheses concern expected mean differences between changes in scores on the instrument in groups or expected correlations between changes in scores on the instrument and changes in scores on other instruments that are known to have adequate responsiveness. One could also consider relative correlations, for example, one may hypothesize that the change on instrument A is expected to correlate more with the change on instrument B than with the change on instrument C because the constructs being measured by instruments A and B are more similar than the construct being measured with instrument C.

Testing hypotheses is much less common in responsiveness studies than in validity studies. However, without specific hypotheses, the risk of bias is high, because retrospectively it is tempting to think up alternative explanations for low correlations instead of concluding that an instrument may not be responsive. Another advantage of defining explicit hypotheses is that it makes the interpretation of the data more transparent because it enables quantification of the number of correlations of differences that are in accordance with the hypotheses.

The hypotheses for testing responsiveness should include the expected direction (positive or negative) and the (absolute or relative) magnitude of the correlations or differences between the change scores. For example, one may expect a positive correlation of at least 0.50 between changes on two instruments that intend to measure the same construct. Or, one may expect that the change in score on instrument A correlates at least 0.10 points higher with the change in score on instrument B than with the change in score on instrument C (see example below). Without this specification of the expected differences or correlations, it is difficult to decide afterwards whether the hypothesis is confirmed or not.

One should not rely on p-values of the correlations, because it is not relevant to determine whether correlations differ statistically significantly from zero. Instead, the responsiveness issue concerns whether the direction and magnitude of the observed correlation is similar to what was expected, based on the construct that is being measured. When assessing differences between changes in groups, it is also less relevant whether these differences are statistically significant (which partly depends on the sample size) than whether these differences are as large as was hypothesized.

## Some Remarks on Alternative Measures of Responsiveness

Many studies use effect sizes as a measure of responsiveness, calculated as the mean change score in a group of patients divided by the standard deviation of the baseline scores or the SD of the change scores. Authors conclude that their instrument is responsive if the effect size is large. This does not have to be true, because effect sizes are measures of the *magnitude* of the change scores rather than the *validity* of the change scores (which is the definition of responsiveness). A high magnitude of change gives little indication of the ability of the instrument to detect change over time on the construct to be measured because the observed change might be smaller or larger than the true change in ability to transfer (e.g., due to floor or ceiling effects or a lack of content validity). Furthermore, effect sizes are highly dependent on the SD and will therefore be higher in a relatively homogeneous population or if the variation in treatment effect is small. Therefore, effect sizes by itself provide very limited evidence of responsiveness. However, if effect sizes are used in a construct approach with a priori defined hypotheses about the expected magnitude of the effect sizes, then it is an acceptable method.

Some authors use the p-value obtained from a paired t-test as a measure of responsiveness and conclude that their instrument is responsive if the p-value is small. This also does not have to be true, because the p-value is a measure of the statistical *significance* of the change scores instead of the *validity* of the change scores. Statistical significance depends on the magnitude of change, the SD of the change scores, and the sample size. Therefore, the paired t-test should not be considered as a good parameter of responsiveness.

Guyatt et al. introduced a responsiveness ratio, defined as the minimal important change (MIC, which is the smallest change in score that patients consider important) on an instrument, and divided by the SD of change scores in stable patients (Guyatt, Walter, & Norman, 1987). Although this ratio is very useful for interpreting change scores, it gives no information about the validity of the change scores. Therefore, it does not fit within the COSMIN definition of responsiveness (see also discussion below). The numerator (MIC) is a measure of the *interpretability* of the change scores, and not the *validity* of the change scores. The denominator (SD of the change scores in stable patients) is an assessment of measurement error.

### Discussion

The COSMIN group defined responsiveness as "the ability of an instrument to detect change over time in the construct to be measured." According to this definition, responsiveness is an aspect of validity. However, other authors use different definitions and subsequently, different methods. For example, Guyatt et al. make a distinction between responsiveness (which they define as the responsiveness ratio mentioned above) and longitudinal validity, which they define as the validity of change scores (Guyatt et al., 1987, 1989). One should be aware of these differences in the literature. Despite using different terms and methods, most authors conceptually agree on the important quality aspects of measurement instruments.

## Cross-References

- ▶ Effect Size
- ▶ Minimal Clinically Important Difference (MCID)
- ▶ Minimal Important Difference

## References

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. J. (2011). *Measurement in medicine. A practical guide*. Cambridge: Cambridge University Press.

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases, 40*, 171–178.

Guyatt, G. H., Deyo, R. A., Charlson, M., Levine, M. N., & Mitchell, A. (1989). Responsiveness and validity in health status measurement: A clarification. *Journal of Clinical Epidemiology, 42*, 403–408.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). International consensus on taxonomy, terminology, and definitions of measurement properties: Results of the COSMIN study. *Journal of Clinical Epidemiology, 63*, 737–745.

Norman, G. R., Stratford, P., & Regehr, G. (1997). Methodological problems in the retrospective computation of responsiveness to change: The lesson of Cronbach. *Journal of Clinical Epidemiology, 50*, 869–879.

# Restlessness

- ▶ Anxiety

# Restoration of Neighborhoods

- ▶ Gentrification

# Restorative Natural Environments

Marc G. Berman[1] and Katherine M. Krpan[2]
[1]The University of South Carolina, Columbia, SC, USA
[2]Department of Psychology, University of South Carolina, Columbia, SC, USA

## Synonyms

Attention restoration theory; Forest bathing/Shinrin-yoku; Nature restoration

## Definition

The restorative effects of interacting with nature refer to theory and research examining

how and why interacting with natural environments (such as nature parks) can improve human psychological health, physical health, and overall well-being.

## Description

Throughout history, many people have known, instinctively, that being in nature is good for us. For example, Buddhist monks sensed this and performed meditative practice in beautiful outdoor settings. Early urban planners also had this insight and tried to incorporate large parks into urban city centers. While the restorative powers of nature may be intuitive, science can provide answers about *why* interacting with nature may be beneficial and *how* it may benefit us. These are important questions for researchers studying how to optimize the environments that surround us.

Much of the work in this area attempts to identify the various psychological and physiological mechanisms that are altered after a restorative nature experience. This research is timely given the large number of individuals that are living in urban settings, often removed from nature; for the first time in human history, over half of the world's population lives in cities (World Bank, 2012). Research has shown some of the adverse effects of living in urban areas. For example, exposure to urban environments may increase the propensity for ▶ mental illness (Pedersen & Mortensen, 2001; Peen, Schoevers, Beekman, & Dekker, 2010) and social ▶ stress (Lederbogen et al., 2011). This is not to say that living in the city is always bad, or bad for all people. Rather, there may be ways to improve urban environments to curb their potentially taxing qualities, and one such way may be to incorporate more green space into urban areas. Given that recent research has shown that brief interactions with nature can have beneficial effects on cognitive functioning (Berman et al., in press; Berman, Jonides, & Kaplan, 2008; Cimprich & Ronis, 2003; Faber Taylor, & Kuo, 2009; Kaplan & Berman, 2010) and more sustained exposure to nature can provide marked mental and physical health benefits (Kuo & Sullivan, 2001; Mitchell & Popham, 2008; Ulrich, 1984), interacting with nature may lead to improvements in human functioning in urban environments. Therefore, incorporating the natural environment into ▶ urban design and cultivating more interaction with nature in cities may lead to improvements in the way humans function in urban environments.

### Theories of Nature Restoration

One of the first theories to explain why interacting with nature is beneficial for human functioning is Attention Restoration Theory (ART; Berman et al., 2008; Kaplan, 1995, 2001; Kaplan & Berman, 2010). ART is based on research showing the separation of attention into two components: involuntary attention, where attention is captured by inherently intriguing or important stimuli, and voluntary or directed attention, where attention is directed by cognitive-control processes. This distinction between voluntary and involuntary attention was initially proposed by William James (1892) and has subsequently been validated both behaviorally (Jonides, 1981) and neurally (Buschman & Miller, 2007; Fan, McCandliss, Fossella, Flombaum, & Posner, 2005). In addition to top-down control, directed attention involves resolving conflict when one needs to suppress distracting stimulation. ART identifies directed attention as the cognitive mechanism that is restored by interactions with nature. According to ART, interacting with environments rich in inherently fascinating stimuli (e.g., sunsets) evokes involuntary attention modestly, allowing directed attention mechanisms a chance to replenish (Kaplan, 1995). That is, the requirement for directed attention in such environments is minimized, and attention is typically captured in a bottom-up fashion by features of the environment itself. So, the logic is that, after an interaction with the natural environment, one is able to perform better on tasks that depend on directed attention abilities. Unlike natural environments, urban environments contain bottom-up stimulation (e.g., car horns). These stimuli capture attention dramatically, requiring directed attention

to overcome that stimulation (e.g., avoiding traffic and ignoring advertising), and making urban environments less restorative.

Previous research has provided support for the hypothesis that interactions with nature improve attention and memory (Berman et al., 2008; 2012; Berto, 2005; Cimprich & Ronis, 2003; Hartig, Evans, Jamner, Davis, & Gärling, 2003; Herzog, Black, Fountaine, & Knotts, 1997; Nee, Jonides, & Berman, 2007; Ottosson, 2005), thereby providing some support for ART. In fact, Berman et al. (2008) found some of the most direct evidence supporting ART in that interactions with nature improved performance on tasks that demanded directed attention, but not on tasks that were more reliant on involuntary attention.

ART is not the only theory to hypothesize why interacting with nature can be restorative (for a complete review of theories on nature restoration, see Hartig et al., 2011). Briefly, some researchers hypothesize that interacting with nature is beneficial because it improves ▶ mood and feelings of connectedness, thereby leading to cognitive improvements (Mayer, Frantz, Bruehlman-Senecal, & Dolliver, 2009). This theory of connectedness has somewhat similar properties to the biophilia hypothesis which posits that humans have an innate love with all things natural/living (Kellert & Wilson, 1993; Wilson, 1984). Therefore, an interaction with nature may be beneficial because one has an innate drive to be in natural environments. Still others suggest that nature is restorative because it increases ▶ positive affect (van den Berg, Koole, & van der Wulp, 2003). While ART provides more of a cognitive theory as to why nature is restorative, biophilia and connectedness theories provide more of an evolutionary account with a strong affective component.

Still other researchers speculate that interacting with nature is beneficial because it reduces stress levels (Laumann, Gärling, & Stormark, 2003; Ulrich et al., 1991; Valtchanov, Barton, & Ellard, 2010) and can buffer children against stress (Wells & Evans, 2003). More research is required to disentangle these different theories and mechanisms that may drive the restoration effect. From Berman et al. (2008, 2012), it was shown that interacting with nature improves memory and mood, but that the mood gains and the memory gains do not correlate with one another. Therefore, it is possible that interacting with nature may have a multitude of independent salutary effects, but more research in this area is required.

## Applications of Nature Restoration to Physical and Mental Health

While research indicates that interacting with nature can improve memory, attention, mood, and stress levels, it is important to examine if those results extend into clinical domains. Given the incredible economic burden of healthcare, it is important to develop low-cost therapies or supplements to treat existing health problems. Interacting with nature could be a time- and cost-efficient supplement to existing medical interventions.

In one of the first studies to show the potential benefits of nature in a clinical setting, Ulrich (1984) documented the experiences of patients who had a view of nature from their hospital room and found they had better postoperative recovery from gallbladder surgery, compared to patients who had a view of a brick wall. The patients with a nature view had shorter postoperative stays and fewer negative evaluations from nurses and took fewer doses of strong and moderate analgesics (Ulrich, 1984). These results suggest a powerful role of nature in potentially aiding recovery from surgery.

Recently, it has been shown that interacting with nature can have salutary effects for individuals diagnosed with clinical depression, a trademark of which is persistent negative thinking (i.e., rumination). People diagnosed with current major depressive disorder were assigned to walk alone, either in nature or an urban setting, for 50 min. To test if walking exacerbated or reduced depressive thought patterns, participants were prompted to think about a negative autobiographical memory, before going on the walk, to induce depressive rumination. Participants diagnosed with depression showed large memory improvements after the nature walk and also showed large mood improvements (Berman et al., 2012). In fact, the effect size for

the memory improvements was five times larger than that found in previous work in a nonclinical sample (Berman et al., 2008, 2012). While interacting with nature may not replace existing empirically validated therapies for depression, these data are good evidence that interacting with nature may help to supplement/enhance existing treatments for depression.

Short 20-min walks in nature can also have a positive attentional impact on children diagnosed with attention deficit hyperactivity disorder (ADHD; Faber Taylor, & Kuo, 2009). Incredibly, interacting with nature was comparable to the drug Ritalin for treating ADHD symptoms. Moreover, the greener the play environment, the more beneficial the effects of being outdoors. Again, these results suggest that interacting with nature could supplement existing treatments for ADHD.

Interacting with nature is also restorative for women diagnosed with ▶ breast cancer. Women with breast cancer undergoing chemotherapy have reported cognitive dysfunction that has been known as "chemo-brain." Unfortunately, the negative effects of chemotherapy may begin even before treatment (Berman et al., 2012a; Cimprich et al., 2010). Brief interactions with nature three times a week helped to counteract the negative effects that are associated with worry related to breast cancer and also the negative effects of chemotherapy (Cimprich & Ronis, 2003). Women in an experimental group, instructed to interact with nature three times a week, went back to work sooner, reported higher quality of life, and scored higher on memory and attention measures compared to breast cancer patients who were instructed to engage in other non-nature related relaxation activities (e.g., puzzles). This is yet another example of how interacting with nature may serve as a supplement to existing medical interventions.

In summary, there is clear evidence that interacting with nature can have clinically meaningful benefits for those afflicted with a number of varying medical conditions. Thus, interacting with nature could, and should, be considered as a supplement to validated medical treatments. It is time- and cost-effective, and perhaps even more importantly, it is often pleasurable, which greatly increases the probability of patient compliance. In addition to the research mentioned above, there are likely a whole host of other medical conditions that could also be treated by interacting with nature. Empirically, for example, we know that people with traumatic brain injury report feeling better after interacting with nature. People with other disabilities could also benefit from interacting with nature. It is necessary to further examine what clinical groups might benefit from interacting from nature and then to disseminate this information to healthcare providers so that it will reach the patients in need. It is also necessary to investigate the dose–response curve. How much time does one have to spend in nature to see cognitive and affective gains? How long do those gains last? Furthermore, how might we make interacting with nature accessible to patients with little access? Advances in technology will allow us to answer some of these questions.

## Future Directions

In the recent years, there has been a great increase in the number of studies that have examined the potential benefits of interacting with nature. There have also been some studies employing novel experimental techniques, such as virtual reality to examine the restorative effects of nature (Valtchanov et al., 2010). Research using functional magnetic resonance imaging (fMRI) is also underway to understand the neural changes that accompany interactions with nature. Other research has begun to examine the effects of nature in large-scale population studies, to uncover the sustained health effects of living in proximity to green space (Mitchell & Popham, 2008). In future lines of research, it is important to combine both objective and subjective outcome measures. In addition, incorporating cutting-edge tools from psychology, urban planning, neuroscience, ecology, physiology, and statistics will help to push the field forward and increase our understanding of how and why nature may benefit us.

## Conclusion

It appears that nature can lower stress levels, improve mood, and also improve memory and attention performance. Interacting with nature

can speed recovery from surgery, benefit women with breast cancer, improve ADHD symptoms in children, and enhance mood and memory for those diagnosed with depression. Therefore, it appears that nature's impact is wide reaching. Given these positive findings, it is also important to provide a note of caution. While we believe that interacting with nature has incredible power to improve human health and well-being, nature is not a panacea capable of curing all ills. We need to find the boundary conditions of nature's effect. It is possible that some human processes may get worse after an interaction with nature and nature may have no impact on certain maladies. With more research in this area, we will be able to define these bounds. In addition, nature certainly is not the only type of environment that could be restorative; some research suggests that places of worship and meditation may also be restorative (Herzog, Ouellette, Rolens, & Koenigs, 2010). The future is certainly very bright, and as more research is performed in this area with a myriad of experimental techniques, we will gain a more nuanced perspective of nature's power to heal and improve human well-being.

## Cross-References

- Affective Connection to Nature
- Attention-Deficit/Hyperactivity Disorder (ADHD)
- Breast Cancer
- Cognitive Function
- Depression in Middle Age
- Healthcare Expenditures as a Percentage of the GNP
- Mental Illness
- Mood
- Nature and Well-being
- Nature Relatedness and Subjective Well-Being
- Physical Quality of Life
- Positive Affect
- Religion, Psychological Well-Being, and Health
- Stress
- Urban Design

## References


Berman, M. G., Askren, M. K., Jung, M. S., Clark, P. M., Therrien, B., Peltier, S., et al. (in press). Chemo brain may not be all about chemotherapy: Pretreatment worry and neurocognitive responses in women with breast cancer. Health Psychology.

Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological Science, 19*(12), 1207.

Berman, M. G., Kross, E., Krpan, K. M., Askren, M. K., Burson, A., Deldin, P. J., et al. (2012). Interacting with nature improves cognition and affect for individuals with depression. *Journal of Affective Disorders, 140*(3), 300–305.

Berto, R. (2005). Exposure to restorative environments helps restore attentional capacity. *Journal of Environmental Psychology, 25*(3), 249.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science, 315*(5820), 1860–1862.

Cimprich, B., Reuter-Lorenz, P., Nelson, J., Clark, P. M., Therrien, B., Normolle, D., et al. (2010). Prechemotherapy alterations in brain function in women with breast cancer. *Journal of Clinical and Experimental Neuropsychology, 32*(3), 324–331.

Cimprich, B., & Ronis, D. L. (2003). An environmental intervention to restore attention in women with newly diagnosed breast cancer. *Cancer Nursing, 26*(4), 284.

Faber Taylor, A. F., & Kuo, F. E. (2009). Children with attention deficits concentrate better after walk in the park. *Journal of Attention Disorders, 12*(5), 402–409.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26*(2), 471–479.

Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S., & Gärling, T. (2003). Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology, 23*(2), 109–123.

Hartig, T., van den Berg, A. E., Hagerhall, C. M., Tomalak, M., Bauer, N., Hansmann, R., et al. (2011). Health benefits of nature experience: Psychological, social and cultural processes. In K. Nilsson, M. Sangster, C. Gallis, T. Hartig, S. De Vries, K. Seeland, & J. Schipperijn (Eds.), *Forest, trees and human health* (pp. 127–168). Dordrecht: Springer Science Business and Media. doi:10.1007/978-90-481-9806-1-5.

Herzog, T. R., Black, A. M., Fountaine, K. A., & Knotts, D. J. (1997). Reflection and attentional recovery as distinctive benefits of restorative environments. *Journal of Environmental Psychology, 17*(2), 165–170.

Herzog, T. R., Ouellette, P., Rolens, J. R., & Koenigs, A. M. (2010). Houses of worship as restorative environments. *Environment and Behavior, 42*(4), 395–419.

James, W. (1892). *Psychology: The briefer course*. New York: Holt.

Jonides, J. (Ed.). (1981). *Voluntary vs. automatic control over the mind's eye's movement*. Hillsdale, NJ: Erlbaum.
Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*(3), 169–182.
Kaplan, S. (2001). Meditation, restoration, and the management of mental fatigue. *Environment and Behavior, 33*(4), 480.
Kaplan, S., & Berman, M. G. (2010). Directed attention as a common resource for executive functioning and self-regulation. *Perspectives on Psychological Science, 5*(1), 43–57.
Kellert, S. R., & Wilson, E. O. (1993). *The Biophilia hypothesis*. Washington, DC: Island Press.
Kuo, F. E., & Sullivan, W. C. (2001). Environment and crime in the inner city: Does vegetation reduce crime? *Environment and Behavior, 33*(3), 343–367.
Laumann, K., Gärling, T., & Stormark, K. M. (2003). Selective attention and heart rate responses to natural and urban environments. *Journal of Environmental Psychology, 23*(2), 125–134.
Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., et al. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature, 474*(7352), 498–501. doi:10.1038/nature10190.
Mayer, F. S., Frantz, C. M., Bruehlman-Senecal, E., & Dolliver, K. (2009). Why is nature beneficial? The role of connectedness to nature. *Environment and Behavior, 41*(5), 607–643.
Mitchell, R., & Popham, F. (2008). Effect of exposure to natural environment on health inequalities: An observational population study. *Lancet, 372*(9650), 1655–1660.
Nee, D. E., Jonides, J., & Berman, M. G. (2007). Neural mechanisms of proactive interference-resolution. *NeuroImage, 38*, 740–751.
Ottosson, J. G. P. (2005). A comparison of leisure time spent in a garden with leisure time spent indoors: On measures of restoration in residents in geriatric care. *Landscape Research, 30*, 23–55.
Pedersen, C. B., & Mortensen, P. B. (2001). Evidence of a dose–response relationship between urbanicity during upbringing and schizophrenia risk. *Archives of General Psychiatry, 58*(11), 1039–1046.
Peen, J., Schoevers, R. A., Beekman, A. T., & Dekker, J. (2010). The current status of urban–rural differences in psychiatric disorders. *Acta Psychiatrica Scandinavica, 121*(2), 84–93.
Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science, 224*(4647), 420–421.
Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology, 11*(3), 201–230.
Valtchanov, D., Barton, K. R., & Ellard, C. (2010). Restorative effects of virtual nature settings. *Cyberpsychology, Behavior and Social Networking, 13*(5), 503–512.
van den Berg, A. E., Koole, S. L., & van der Wulp, N. Y. (2003). Environmental preference and restoration: (How) are they related. *Journal of Environmental Psychology, 23*(2), 135–146.
Wells, N. M., & Evans, G. W. (2003). Nearby nature: A buffer of life stress among rural children. *Environment and Behavior, 35*(3), 311–330.
Wilson, E. O. (1984). *Biophilia*. Cambridge, MA: Harvard University Press.
World Bank. (2012). *Urban development data*. Retrieved from Oct 2012 http://data.worldbank.org/topic/urban-development

## Retirement Effects, Quality of Life

Nadia Steiber
Department Socioeconomics, Institute for Sociology and Social Research, Vienna University of Economics, Vienna, Austria

### Synonyms

Labor market exit; Transition to retirement; Withdrawal from work; Work-retirement transition

### Definition

This entry is concerned with the impact of the transition from work to *retirement* on the QoL of individuals. Retirement in this entry refers to the commencement of the stage in a person's life course in which he/she is no longer gainfully employed (*full retirement*), although *gradual retirement* (involving work hour reductions and phasing into retirement) will also be considered. Moreover, a differentiation is made between *regular retirement* at the standard age for retirement and *early retirement*. The *eligibility ages for public old-age benefits* vary across countries and in some countries also between women and men. *Pensionable age* is defined by the OECD (2011) as the age at which people can first draw full benefits (i.e., without actuarial reduction for

**Retirement Effects, Quality of Life, Table 1** Pensionable ages in a select set of OECD countries, 2010

| | Men | Women (if different) |
|---|---|---|
| Australia | 65 | 62 |
| Austria | 65 | 60 |
| Belgium | 60 | |
| Canada | 65 | |
| Czech Republic | 61 | 58.7 |
| Denmark | 65 | |
| Finland | 65 | |
| France | 60.5 | |
| Germany | 65 | |
| Greece | 57 | |
| Hungary | 60 | 59 |
| Ireland | 65 | |
| Italy | 59 | |
| Japan | 64 | 62 |
| Netherlands | 65 | |
| Norway | 67 | |
| Poland | 65 | 60 |
| Portugal | 65 | |
| Slovak Republic | 62 | 57 |
| Spain | 65 | |
| Sweden | 65 | |
| Switzerland | 65 | 63 |
| United Kingdom | 65 | 60 |
| United States | 66 | |

Source: OECD Pensions at a Glance 2011

early retirement). In many countries, it is identical to the *normal pension* age set out in legislation. In others, it is possible to retire earlier than the normal age without an *actuarial* reduction in pension benefits (given a long duration of benefit payment and fulfillment of contribution requirements). Finally, some countries do not have a normal pension age but define a corridor (i.e., a range of ages) at which the pension may first be drawn. *Pensionable ages* as defined by the OECD (2011) are designed to be comparable between countries (see Table 1 below for overview).

## Description

The question of how *retirement* affects the ▶ well-being of older adults, their physical and mental ▶ health, and ▶ quality of life (QoL) as well as the related question of whether or not prolonged employment is conducive to the QoL of older adults is highly topical. The policy discourse at national and international level has put strong emphasis on the issue of *population aging* and its impact on welfare systems in recent years. When birth rates decline while ▶ life expectancy increases, old-age dependency ratios rise unless the number of economically active members of society can be substantially increased. For this reason, a central aim across the OECD and other countries is to raise the employment rates of older adults (for definition of *older or aging workers*, see corresponding entry in this volume). To this end, more or less far-reaching pension reforms have been implemented, aimed at restricting access to and reducing the economic incentives for *early retirement* (Arza & Kohli, 2008).

A large body of empirical research evaluates the *causal effects* of such *reforms* (i.e., mainly retrenchment) on older workers' retirement plans, expectations, and behaviors (e.g., Krueger & Pischke, 1992; Coile & Gruber, 2000, 2007; Chan & Stevens, 2004; Gruber & Wise, 2004; Bottazzi, Jappelli, & Padula, 2006; Belloni & Alessie, 2009; De Grip, Lindeboom, & Montizaan, 2009; Hanel & Riphahn, 2009; Mastrobuoni, 2009; Hanel, 2010). Many of these studies present evidence suggesting that individuals are responsive to the financial incentives embedded in pension plans and social security. Yet, some of the estimated effects of financial incentives may be overstated due to the difficulty of controlling for *unobserved heterogeneity* in studies of retirement behavior: Financial incentive variables and individuals' work orientation (see also ▶ Employment Commitment) tend to correlate (see, e.g., Chan & Stevens, 2004; Filer & Honig, 2005; Hanel & Riphahn, 2009).

Continued work until higher ages is important for securing the sustainability of welfare systems. And some commentators suggest that it may also be conducive to the well-being and overall QoL of older adults. The results from cross-sectional studies (i.e., using data from a single point

in time) tend to suggest that there is a positive correlation between working at older ages (late-life paid work) and individual well-being (for overview, see, e.g., Calvo, 2006). Such findings are usually interpreted with view to the *latent functions of paid work* (Jahoda, 1982) suggesting that working (at any age) has a high psychosocial value over and above the associated monetary rewards. In this tradition, working longer is argued to be conducive to older persons' well-being because it helps maintain a sense of purpose in life, opportunities for the utilization of skills, valuable social contacts at work, and a regular time structure (e.g., Freedman, 2007). Conversely, it has been argued that early retirement may have negative effects on individuals' QoL, their living standard, their mental well-being, health, and ▶ life expectancy. Yet, due to processes of selection and attrition – workers self-selecting into retirement, older adults being sorted out of employment, and individual variation in life expectancy – the evaluation of *causal effects* of retirement on individuals' QoL involves a number of methodological challenges. Depending on the pattern of selection, we may observe *happy older worker* effects or *happy retiree* effects that reflect reverse (bidirectional) causality (see, e.g., Ekerdt, 2010).

For these and other reasons, the empirical evidence on the causal impact of the work-retirement transition on the QoL of individuals is somewhat inconclusive. In the remainder of this entry, some recent evidence from studies using longitudinal data and paying attention to the issues of selection are reviewed.

Calvo, Haverstick, and Sass (2009) study the factors that affect the ▶ happiness of individuals who make the transition from full employment to full retirement (using longitudinal data from the US Health and Retirement Study). They do not find support for the frequent claim (by policy makers) that workers would be happier with a gradual retirement (i.e., phasing out based on part-time work) than with an abrupt end of their careers (i.e., direct transition from full-time work to full retirement, *cold turkey*). Instead their results suggest that it is the degree to which retirement is perceived as voluntary (chosen) or involuntary (forced) that significantly affects retirees' ▶ happiness (see also Calvo, 2006; De Vaus, Wells, Kendig, & Quine, 2007). Investigating ▶ health instead of happiness as an outcome, however, some studies do find that compared to abrupt retirement, gradual retirement has beneficial effects (e.g., De Vaus et al., 2007; Calvo & Sarkisian, 2011).

Calvo and Sarkisian (2011) study the impact of retirement on four dimensions of well-being: individuals' mood, health, income, and social ties. They hypothesize that the impact of the work-retirement transition on the well-being of individuals depends on the *speed of the transition* (gradual/abrupt), on *perceived control* (voluntary/forced), and on the degree to which the transition could be anticipated by the individual (for summary of results, see Table 2 below). Individuals on the gradual retirement path report significantly smaller decreases in ▶ health both at the onset of the transition (immediate effect) and at full retirement (more long-term effect). In other words, a slower phasing into retirement tends to be associated with smaller declines in ▶ health upon retirement. Moreover, those who retire gradually are better off in terms of their income but only during the transition period. Yet, gradual retirees tend to report worse outcomes in terms of their ▶ social well-being, at least at the onset of the transition (i.e., when they are still working). When retirement was chosen by the individual (voluntary rather than forced), this tends to be associated with more favorable changes in mood, health, and income – both in phased and full retirement. When retirement was unexpected, transitions are more likely to have a negative impact on ▶ social well-being. Later retirement (i.e., at a higher age) is associated with more favorable developments in terms of mood and income. Yet, continued work until higher ages appears to come along with greater health problems when compared to earlier retirement. In other words, in this study, retiring later is found to have better psychological and economic outcomes but worse health outcomes (Table 2).

Using longitudinal data from the US Health and Retirement Study, Calvo, Sarkisian, and Tamborini (2011) explore the causal impact of

**Retirement Effects, Quality of Life, Table 2** The effects of retirement on four dimensions of individual well-being, depending on the characteristics of the work-retirement transition

| | | Mood | Health | Income | Social ties |
|---|---|---|---|---|---|
| At the onset of the transition (*immediate effects*) | Gradual retirement | – | Pos | Pos | Neg |
| | Chosen/voluntary retirement | Pos | Pos | Pos | – |
| | Unexpected | – | – | – | – |
| | Retirement age | Pos | Neg | Pos | – |
| At full retirement (*longer-term effects*) | Length of transition | – | Pos | – | – |
| | Chosen/voluntary retirement | Pos | Pos | Pos | – |
| | Unexpected | – | – | – | Neg |
| | Retirement age | Pos | Neg | Pos | – |

Notes: The table summarizes the results from the study by Calvo and Sarkisian (2011) - author's own summary. Pos (neg) signifies a larger increase (decrease) or smaller decrease (increase) in the corresponding well-being measure, depending on the general direction of change (which is a decline in mood, health, and income and no significant change in social ties at the onset of the transition for the references person – a white male with high SES who has been forced into abrupt retirement). At full retirement, the base line is a decline in health and income and no significant change in income and social ties for the reference person.

the timing of retirement on the change in subjective well-being (comparing individuals' subjective health and mood before and after the transition). They find that ▶ subjective well-being is highest when retirement happens *on time*, i.e., neither very early nor very late. In the current US context, the optimal retirement transition is estimated to start around age 62. For this reason, the authors suggest that the rise in pensionable ages will probably reduce well-being (at least in the short run). Using data from the *English Longitudinal Study of Ageing* (ELSA), Bound and Waidman (2007) find a slightly positive health effect of retirement at the *normal retirement age* once the possibility of endogenous entry into retirement is accounted for. Using the Survey of Health, Ageing and Retirement in Europe (SHARE), also Coe and Zamarro (2011) find retirement to have health-preserving effects (see also Neuman, 2008). Other studies find negative effects of *early retirement* (using instrumental variables approaches). Rohwedder and Willis (2010), for instance, find detrimental effects of *early retirement* on cognitive ability. Kuhn, Wuellrich, and Zweimüller (2010) find that early exit from the labor force increases men's mortality (but not that of women). They estimate that one additional year of early retirement increases the probability of dying before the age of 67 by 2.4 percentage points (from an average probability of 18 percent). Potential reasons for this finding include changes in lifestyle upon retirement with negative health implications.

Overall, the question of how retirement transitions affect individuals' well-being and QoL is still unresolved. The most recent studies tend to suggest that a health risk can be inherent both in continued work until very high ages as well as in (too early) retirement. The effects of work-retirement transitions crucially depend on the *timing* and the *pace* of the transition and the degree to which individuals perceive *control* over this process. Hence, a greater flexibility in retirement options in terms of voluntary part-time work, temporary retirement, etc., may not only help slowing down the retirement process (phased retirement) but also increase the degree to which the worker perceives having some control over this process – with favorable outcomes for individual well-being (e.g., Zhan, Wang, Liu, & Schultz, 2009, showing health benefits of engaging in *bridge employment* for retirees in the USA, using the first four waves of the Health and Retirement Study).

## Cross-References

▶ Aging Workers and the Quality of Life
▶ Birth Rate
▶ Employment Commitment

- ▶ Happiness
- ▶ Life Expectancy
- ▶ Social Well-being
- ▶ Subjective Well-being

## References


Arza, C., & Kohli, M. (Eds.). (2008). *Pension reform in Europe: Politics, policies and outcomes*. London/New York: Routledge.

Belloni, M., & Alessie, R. (2009). The importance of financial incentives on retirement choices: New evidence for Italy. *Labour Economics, 16*, 578–588.

Bottazzi, R., Jappelli, T., & Padula, M. (2006). Retirement expectations, pension reforms, and their impact on private wealth accumulation. *Journal of Public Economics, 90*, 2187–2212.

Bound, J. & Waimann, T. (2007). *Estimating the health effects of retirement*. University of Michigan Retirement Research Center. WP 2007–168.

Calvo, E. (2006). *Does working longer make people healthier and happier?* MPRA Paper No. 5606, Munich Personal RePEc Archive. http://mpra.ub.uni-muenchen.de/5606/1/MPRA_paper_5606.pdf. Retrieved February 12, 2012.

Calvo, E., Haverstick, K., & Sass, S. A. (2009). Gradual retirement, sense of control, and retirees' happiness. *Research on Aging, 31*, 112–135.

Calvo, E., & Sarkisian, N. (2011). *Retirement and well-being: Examining the characteristics of life course transitions*. Working Paper #2, Institute of PublicPolicy at Universidad Diego Portales, Santiago,Chile.http://www.politicaspublicas.udp.cl/media/publicaciones/archivos/263/Documento_Completo.pdf. Retrieved February 12, 2012.

Calvo, E., & Sarkisian, N., & Tamborini, C. R. (2011). *Causal effects of retirement timing on subjective well-being: The role of cultural norms and institutional policies*. Working Paper #17, School of Business and Economics at Universidad Diego Portales, Santiago, Chile.

Chan, S., & Stevens, A. H. (2004). Do changes in pension incentives affect retirement? A longitudinal study of subjective retirement expectations. *Journal of Public Economics, 88*, 1307–1333.

Coe, N. B., & Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics, 30*, 77–86.

Coile, C., & Gruber, J. (2000). *Social security and retirement*. National Bureau of Economic Research (NBER) Working Paper No. 7830.

Coile, C., & Gruber, J. (2007). Future social security entitlements and the retirement decision. *The Review of Economics and Statistics, 89*, 234–246.

De Grip, A., Lindeboom, M., & Montizaan, R. (2009). *Shattered dreams: The effects of changing the pension system. Late in the Game*. IZA Discussion Paper No. 4034. http://ftp.iza.org/dp4034.pdf. Retrieved March 21, 2012.

De Vaus, D., Wells, Y., Kendig, H., & Quine, S. (2007). Does gradual retirement have better outcomes than abrupt retirement? Results from an Australian panel study. *Ageing and Society, 27*, 667–682.

Ekerdt, D. J. (2010). Frontiers of research on work and retirement. *Journal of Gerontology: Social Sciences, 65B*, 69–80.

Filer, R.K., & Honig, M. (2005). Endogenous pensions and retirement behavior. CESIFO Working Paper No. 1547. http://www.cesifo-group.de/portal/pls/portal/docs/1/1188540.PDF. Retrieved March 21, 2012.

Freedman, M. (2007). *Encore: Finding work that matters in the second half of life*. New York: Public Affairs.

Gruber, J., & Wise, D. (2004). *Social Security programs and retirement around the world: Micro-estimation*. Chicago: University of Chicago Press/National Bureau of Economic Research (NBER).

Hanel, B. (2010). Financial incentives to postpone retirement and further effects on employment – Evidence from a natural experiment. *Labour Economics, 17*, 474–486.

Hanel, B., & Riphahn, R. T. (2009). *New evidence on financial incentives and the timing of retirement*. BGPE Discussion Paper No. 76. http://www.bgpe.de/texte/DP/076_hanel.pdf. Retrieved March 21, 2012.

Jahoda, M. (1982). *Employment and unemployment. A social-psychological analysis*. Cambridge, UK: Cambridge University Press.

Krueger, A., & Pischke, J.-S. (1992). The effect of social security on labor supply: A cohort analysis of the notch generation. *Journal of Labor Economics, 10*, 412–437.

Kuhn, A., Wuellrich, J.-P., & Zweimüller, J. (2010). *Fatal attraction? Access to early retirement and mortality*. IZA DP No. 5160. http://ftp.iza.org/dp5160.pdf. Retrieved May 1, 2012.

Mastrobuoni, G. (2009). Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities. *Journal of Public Economics, 93*, 1224–1233.

Neuman, K. (2008). Quit your job and get healthier? The effect of retirement on health. *Journal of Labor Research, 29*, 177–201.

OECD. (2011). *Pensions at a Glance 2011*. Paris: Author.

Rohwedder, S., & Willis, R. J. (2010). Mental retirement. *Journal of Economic Perspectives, 24*, 119–138.

Zhan, Y., Wang, M., Liu, S., & Schultz, K. S. (2009). Bridge employment and retirees' health: A longitudinal investigation. *Journal of Occupational Health Psychology, 14*, 374–389.


## Retirement Planning and QWL

- ▶ Work Life and Retirement Planning, Quality of

# Retirement, Emotional Well-Being, and Work-Family Conflict

Jack Lam[1] and Kathryn M. Coursolle[2]
[1]Sociology Department and Minnesota Population Center, University of Minnesota, Minneapolis, MN, USA
[2]Minnesota Population Center, University of Minnesota, Twin Cities, Minneapolis, MN, USA

## Definition

► Work-family conflict is a type of interrole conflict that occurs when the demands of work and family are perceived as incompatible (Greenhaus & Beutell, 1985). Retirement is a life stage that traditionally implies a full exit from the workplace and has been understood to be for ► enjoyment and full-time leisure. However, in reality, the experience of retirement may vary across individuals. Workers of all life stages may face some form of work-family conflict, as they confront the demands of paid work with caring for spouse, children, parents, and kin, but researchers focus less attention on these issues later in the life course. Rising life expectancies and the aging of the baby boom cohort highlight the importance of understanding the relationship between retirement and well-being in later life. Work-family conflict may be a particularly salient component for investigating postretirement well-being, since it may influence retirement timing, ► motivation, and the postretirement context. This entry reviews the current research on retirement, ► emotional well-being, and work-family conflict and suggests directions for future research on this topic.

## Description

Although little work has examined the role of work-family conflict in predicting ► psychological well-being after retirement to date, several studies find that characteristics of the prior family and work environment are associated with postretirement emotional and/or physical well-being. For instance, in terms of family context, researchers find that retired men who have a wife that is currently employed report greater depressive symptoms than retired men with wives who are not working (Szinovacz & Davey, 2004). Research also suggests that several aspects of one's preretirement work environment or subjective assessments of one's job (positive or negative) are related to ► emotional well-being upon retirement. For example, Kubicek, Korunka, Raymo, and Hoonakker (2011) find that high levels of ► job satisfaction prior to retirement is associated with lower positive psychological functioning postretirement for men. For women, higher wages and viewing one's work as important is associated with poorer depressive symptoms upon retirement. Prior work also found that whereas high employment grade is related to better mental function after retirement (Mein, Martikainen, Hemingway, Stansfeld, & Marmot, 2003), poor work environments are associated with improved self-rated health after retirement (Westerlund et al., 2009). Until recently, however, little was known about how prior stress originating from the interface between work and family roles may shape postretirement ► emotional well-being.

One exception is a study by Coursolle, Sweeney, Raymo, and Ho (2010). This paper investigates whether and how prior exposure to ► work-family conflict in late midlife moderates the association between retirement and ► emotional well-being. They utilize Wheaton's (1990) contextual model of role transitions which suggests that there may be an interaction between life transitions and prior role stress. For example, the effects of retirement on emotional well-being may vary depending on the level of prior stress experienced in the employee role. The authors hypothesize that retirement may be associated with relatively better ► emotional well-being outcomes for individuals who previously experienced high levels of ► work-to-family conflict. For these individuals, retirement may provide a source of relief since work-based sources of stress should be ameliorated upon retirement. Their hypotheses,

however, for individuals experiencing high levels of ▶ family-to-work conflict are more uncertain. They suggest that on the one hand retirement may ease the active struggle between work and family possibly improving family conditions. In this way, retirement may have positive implications for emotional well-being. On the other hand, retirement may eliminate a source of ▶ social support, and the pressures associated with ▶ caregiving, ▶ housework, and marital issues may persist after retirement, especially for women. Thus, retirement for these individuals could also be associated with relatively poorer emotional well-being.

The essay makes use of a unique dataset, the Wisconsin Longitudinal Study (WLS), which follows a random sample of 10,317 Wisconsin high school graduates from the class of 1957. The graduates were interviewed at several points in time between 1957 and 2004. The analysis focuses on the 1993 and 2004 waves of the WLS. The sample allows for an examination of postretirement ▶ emotional well-being when the graduates were between 64 and 65 years old in 2004 while controlling for preretirement levels of ▶ psychological well-being and ▶ work-family conflict assessed in 1993. The final analytical sample is restricted to individuals who were continuously married between the 1993 and 2004 waves, had not yet retired in 1993, and had information on variables included in the analysis. The final analytic sample includes 2,666 cases for the analysis of depressive symptoms and 2,855 individuals for the analysis of positive psychological functioning.

The measure of depressive symptoms is the ▶ Center for Epidemiologic Studies Depression (CES-D) scale, which consists of 20 items asking respondents how many days in the past week they experienced a particular depressive symptom (e.g., "I felt depressed," "I had crying spells," and "I felt lonely") (Radloff, 1977). The second measure is an abridged version of Ryff's (1989) scale of ▶ psychological well-being, which consists of 20 items assessing self-acceptance, positive relations with others, autonomy, environmental mastery, purpose in life, and ▶ personal growth. Since the source of ▶ work-family conflict can originate from either the family or work environment, two measures are used. One measure is a scale assessing perceived work-to-family conflict and the other measures perceived ▶ family-to-work conflict. Perceived retirement indicates whether the respondent was not at all retired, partially retired, or fully retired in 2004. In addition, factors that may be related to retirement or emotional well-being are incorporated as control variables such as marital quality, spouse's labor force status, respondent and spouse self-rated health, caregiving responsibilities, whether children were living in the household, characteristics of the current or prior job, and basic demographic variables.

Coursolle and colleagues employ two types of statistical models to investigate the relationship between retirement, emotional well-being, and prior work-family conflict. First, they use ordinary least squares regression to examine how prior ▶ work-family conflict moderates the association between retirement and psychological well-being. Each outcome measure from 2004 is regressed on the same emotional well-being measure from 1993, a variable indicating whether the respondent retired between the waves, perceived ▶ work-family conflict assessed in 1993, an interaction term between ▶ work-family conflict and retirement status, and the control variables. However, if certain personality or background traits (e.g., tendency towards good organization) are associated with the experience of retirement or ▶ work-family conflict and are not controlled for in the analysis, the conventional regression estimates may be biased. Therefore, in order to examine the robustness of findings to control for unobserved characteristics that are constant over time, fixed-effects models are utilized in the second stage of analysis. The authors also investigate how the results from both sets of analysis differ by sex.

The results from the ordinary least squares models suggest that prior ▶ work-family conflict moderates the association between retirement and ▶ emotional well-being. However, findings differ depending on whether the ▶ work-family conflict originates from the family or work environments and which outcome measure is

examined. For example, retirement (partial and full) is associated with lower levels of depressive symptoms among men and women who experienced high levels of prior ▶ work-to-family conflict. Retirement for this group may represent more of a relief than a stressor. However, with regard to ▶ family-to-work conflict, the authors find different results by gender. For men with high levels of prior ▶ family-to-work conflict, partial and full retirement is associated with lower depressive symptoms, but the reverse is true for women. Women with high levels of prior ▶ family-to-work conflict but subsequently retired report greater depressive symptoms than women who did not enter retirement. This suggests that women with stress originating in the family may not gain relief from retirement. Indeed, retirement may not eliminate stressful family demands and could eliminate a potential source of ▶ social support from the workplace. These factors may be felt more strongly by women than men (Hochschild, 1997). In regard to the other outcome measure, there was no evidence that ▶ work-family conflict moderates the association between retirement and positive psychological functioning.

Next the authors estimate fixed-effects models to determine whether the findings from the linear models persist after adjusting for unobserved factors that are constant over time. The results from the analysis of depressive symptoms are generally consistent with the results from the ordinary least squares models. With regard to positive psychological functioning, the authors find somewhat different results in the fixed-effects models compared to the first stage of analysis. Unlike the findings from the ordinary least squares models, once adjusting for unobserved time-invariant variables, suggestive evidence emerges that ▶ work-family conflict may moderate the relationship between retirement and positive psychological functioning. The authors suggest that this may occur because the fixed-effects models control for unobserved personality or coping style differences, such as time management or goal directedness, that may also be associated with aspects of Ryff's positive well-being scale.

Coursolle and colleagues' study highlights a number of ways ▶ work-family conflict is important for better understanding the relationship between retirement and ▶ emotional well-being. However, this essay has limitations that suggest opportunities for future research on the topic. For example, measures of preretirement work and family context in this study may not fully capture Wheaton's concept of chronic prior role stress since they are all gathered from one point in time (the 1993 wave). Future work should investigate the implications of variation in measuring prior role stress, as well as capturing a longer duration of time. In addition, findings from this study are based on a sample of mostly White high school graduates from Wisconsin, and may not generalize to the broader US population of similar ages. The WLS also represents the experiences of the leading edge of the baby boom. Cohorts since then have experienced dramatic shifts in the work and family realms. For one, the workplace has undergone many changes in recent years, including global competition, and technological advances, leading workers to report heightened ▶ job insecurity (Burchell, 2002). Employees may now face pressure to work longer hours in order to portray themselves as efficient and productive workers and this may have important implications for ▶ work-family conflict (Blair-Loy, 2003; Jacobs & Gerson, 2004). Labor force demographics have also changed, with higher female participation and a rise in dual-earner households. In addition, gains in ▶ life expectancy means that workers are now more likely than previous generations to have to combine paid employment with the demands of caring for elderly parents, a spouse in poor health, and grandchildren. Lastly, retirement is increasingly a process, rather than a single transition, as workers engage in phased retirement, or become involved in ▶ volunteering, or postretirement employment, necessitating a broader understanding and definition of retirement (Moen, 2003). These demographic trends, combined with the changing nature of employment and retirement, provide impetus to better understand the work-family interface confronting future retirees.

## References


Blair-Loy, M. (2003). *Competing devotions: Career and family among women executives*. Cambridge, MA: Harvard University Press.

Burchell, B. (2002). The prevalence and redistribution of job insecurity and work intensification. In B. Burchell, D. Ladipo, & F. Wilkinson (Eds.), *Job insecurity and work intensification* (pp. 61–76). London: Routledge.

Coursolle, K., Sweeney, M., Raymo, J., & Ho, J.-H. (2010). The association between retirement and emotional well-being: Does prior work-family conflict matter? *Journal of Gerontology: Social Sciences, 65B*(5), 609–620.

Greenhaus, J. H., & Beutell, N. J. (1985). Sources of conflict between work and family roles. *Academy of Management Review, 10*, 76–88.

Hochschild, A. R. (1997). *The time bind: When work becomes home and home becomes work*. New York: Henry Holt.

Jacobs, J. A., & Gerson, K. (2004). Understanding changes in American working time: A synthesis. In C. F. Epstein & A. L. Kalleberg (Eds.), *Fighting for time: Shifting boundaries of work and social life* (pp. 25–45). New York: Russell Sage.

Kubicek, B., Korunka, C., Raymo, J., & Hoonakker, P. (2011). Psychological well-being in retirement: The effects of personal and gendered contextual resources. *Journal of Occupational Health Psychology, 16*(2), 230–246.

Mein, G., Martikainen, P., Hemingway, H., Stansfeld, S., & Marmot, M. (2003). Is retirement good or bad for mental and physical health functioning? Whitehall II longitudinal study of civil servants. *Journal of Epidemiology and Community Health, 57*, 46–49.

Moen, P. (2003). Midcourse: Navigating retirement and a new life stage. In J. Mortimer & M. Shanahan (Eds.), *Handbook of the life course* (pp. 269–291). New York: Kluwer Academic/Plenum.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*, 1069–1081.

Szinovacz, M. E., & Davey, A. (2004). Honeymoons and joint lunches: Effects of retirement and spouse's employment on depressive symptoms. *Journal of Gerontology: Psychological Sciences, 59B*, P233–P245.

Wheaton B. (1990). Life transitions, role histories, and mental health. *American Sociological Review, 55*, 209–223.

Westerlund, H., Kivimaki, M., Singh-Manoux, A., Melchior, M., Ferrie, J., Pentti, J., et al. (2009). Self-rated health before and after retirement in France (GAZEL): A cohort study. *Lancet, 374*, 1889–1896.


# Retrospective Reports


Michael D. Robinson
Department of Psychology, North Dakota State University, Fargo, ND, USA


## Synonyms

Trait reporting

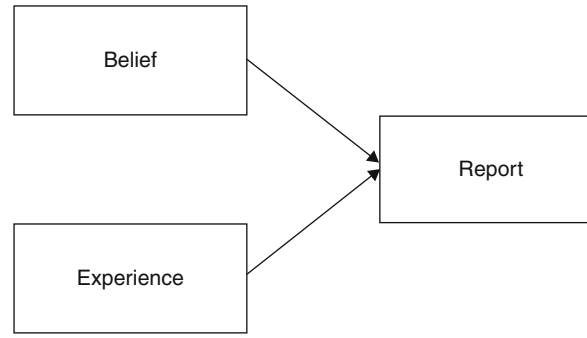
## Definition

Retrospective reports are those in which individuals are asked to characterize their subjective well-being or emotions in the past or in general. Reports of this type are foundational to the personality and clinical literatures. Retrospective reports of life satisfaction, emotion, anxiety, and depression are quite common.

## Description

The cognitive memory literature has long considered retrospective reports potentially suspect (Bartlett, 1932). Memory, from this perspective, is a constructive process that often results in significant distortions of prior events and experiences. Yet, individuals in multiple literatures are commonly asked to report on their emotions, ▶ life satisfaction, or subjective well-being in a retrospective manner. Such reports are very reliable, but likely biased.

Biases in retrospective reports seem to take a general form in which emotion-related beliefs are more predictive of retrospective reports of emotion or subjective well-being than reports based on more momentary experiences (Robinson & Clore, 2002a). For example, women are viewed as the more "emotional" sex and such a pattern is more evident in retrospective reports of emotion than justified on the basis of daily reports of emotion (Barrett, Robin, Pietromonaco, & Eyssell, 1998). Similar biasing effects have been demonstrated for ethnicity (Oishi, 2002) and individual differences




**Retrospective Reports, Fig. 1** Two broad influences on reporting

in the personality trait of neuroticism (Larsen, 1992). Manipulated beliefs also result in biased retrospective reports of emotion (Klaaren, Hodges, & Wilson, 1994). Figure 1 provides a simple, parsimonious model for such dissociations. Reports of emotion or subjective well-being can either be made on the basis of *beliefs* or *experiences* and these are separable influences. With the passage of time, particular experiences are forgotten or too difficult to retrieve, thus shifting retrospective reports in a belief-driven direction.

The simple model of Fig. 1 is supported by neurological studies showing that there are two memory systems in the brain – one that preserves specific moments of experience and one that does not (McClelland, McNaughton, & O'Reilly, 1995). Memory for specific experiences decays in the matter of weeks. On the basis of such neurological considerations, several studies asked individuals to report on their emotional experiences over seven time frames – right now, last few hours, last few days, last few weeks, last few months, last few years, and "in general" (Robinson & Clore, 2002b). Strikingly, reaction time and priming evidence converged on the idea that reports of emotion concerning time frames less than 2 weeks were made by retrieving specific experiences, but this was not the case for time frames greater than 2 weeks. Multiple predictions of this two-process model of reporting were subsequently confirmed (Robinson & Sedikides, 2009). In sum, it appears theoretically important to make a sharp distinction between reports of emotion or subjective well-being that are belief-based versus experience-based.

In the absence of sophisticated cognitive methods, it is difficult to isolate such influences on reporting. Further, a case can be made that beliefs about emotion or subjective well-being are often generally accurate. For example, although individuals high in neuroticism may sometimes retrospectively exaggerate their distress (Larsen, 1992), it is clear that they are generally more distressed than are individuals low in neuroticism (Widiger, Verheul, & van den Brink, 1999). Nonetheless, some beliefs about emotion or subjective well-being may be entirely erroneous (Wilson & Gilbert, 2005). Certainly, we need a better understanding of which beliefs are erroneous, which are largely accurate, and why such beliefs differ in this regard.

Documenting retrospective biases and understanding their nature is important. More important, though, is the implication that retrospective reports of emotion or subjective well-being should not necessarily be taken at face value. Is there another way of understanding the quality of life of individuals? In an influential statement, Mischel (1968) suggested that the trait approach to assessment does not predict state-related outcomes to the extent desirable. As an alternative, it might be better to bypass trait or retrospective reports entirely and simply assess states (i.e., momentary or recent experiences). Although they are quite variable and have many causes, averaging across enough state reports produces reliable knowledge concerning the individual's general subjective well-being (Diener & Larsen, 1984).

Such ideas gave rise to experience-sampling approaches to assessment. By paper-and-pencil,

pre-programmed wrist-watches, palm pilots, cell phones, or internet connections, individuals can be asked to report on their everyday experiences repeatedly over time (Conner, Tennen, Fleeson, & Barrett, 2009). So far, however, such technologies have primarily been used to understand trait functioning. For example, experience-sampling studies have shown that individuals higher in the trait of neuroticism react with greater negative emotion to daily stressors (Suls & Martin, 2005). Studies of this type are important, but are unlikely to supplant traditional approaches to personality assessment, which typically use retrospective or trait instructions and time frames (Robinson & Clore, 2002b).

### Discussion

It is implausible that individuals retrieve specific experiences and aggregate them when making some retrospective self-reports of their emotions or subjective well-being. Rather, they retrieve self-knowledge that is more general and belief-based in making such self-reports (Robinson & Clore, 2002a, b). Such beliefs can be inaccurate or accurate in characterizing actual experiences, but further research is necessary to understand such moderating factors. Suggestions to replace trait with state assessments have not, and likely will not, supplant typical personality assessment procedures. On the other hand, there are multiple benefits to recognizing that trait and state reports are not the same, including greater curiosity concerning the manner in which individuals actually experience their lives.

## References


Barrett, L. F., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are women the 'more emotional' sex? Evidence from emotional experiences in social context. *Cognition and Emotion, 12*, 555–578.

Bartlett, F. C. (1932). *Remembering*. Oxford: University Press.

Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and Personality Psychology Compass, 3*, 292–313.

Diener, E., & Larsen, R. J. (1984). Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. *Journal of Personality and Social Psychology, 47*, 871–883.

Klaaren, K. J., Hodges, S. D., & Wilson, T. D. (1994). The role of affective expectations in subjective experience and decision-making. *Social Cognition, 12*, 77–101.

Larsen, R. J. (1992). Neuroticism and selective encoding and recall of symptoms: Evidence from a combined concurrent-retrospective study. *Journal of Personality and Social Psychology, 62*, 480–488.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*, 419–457.

Mischel, W. (1968). *Personality and assessment*. Hoboken: Wiley.

Oishi, S. (2002). The experiencing and remembering of well-being: A cross-cultural analysis. *Personality and Social Psychology Bulletin, 28*, 1398–1406.

Robinson, M. D., & Clore, G. L. (2002a). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934–960.

Robinson, M. D., & Clore, G. L. (2002b). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology, 83*, 198–215.

Robinson, M. D., & Sedikides, C. (2009). Traits and the self: Toward an integration. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality* (pp. 457–472). New York: Cambridge University Press.

Suls, J., & Martin, R. (2005). The daily life of the garden-variety neurotic: Reactivity, stressor exposure, mood spillover, and maladaptive coping. *Journal of Personality, 73*, 1485–1510.

Widiger, T. A., Verheul, R., & van den Brink, W. (1999). Personality and psychopathology. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 347–366). New York: Guilford Press.

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science, 14*, 131–134.


## Returns to Education and Post-school Investments

▶ Human Capital Models

## Reverse Logistics

▶ Waste Recycling

## RGH

▶ Health-Related Quality of Life and Reliance on God's Help

## Rheumatoid Arthritis

▶ Inflammatory Arthritis, EQ-5D, and SF-6D

## Rhythmic Movement

▶ Dance and the Quality of Life

## Rich People, an Overview

Laura Angelescu
Department of Economics, University of Southern California, Los Angeles, CA, USA

### Synonyms

Wealthy people

### Definition

Rich people are those who have large amounts of money and material goods relative to other people. Within a country or region, the income needed for people to be considered rich varies with the average level of wealth in that country or region.

See also ▶ Affluence.

### Description

When asked what would improve their quality of life, most people answer "more money" (e.g., Campbell, 1981). Based on the assumption that increased ▶ affluence leads to increased happiness, people believe that "pleasure and material comforts should be grasped wherever they can, and that these alone will improve the quality of one's life" (Csikszentmihalyi, 1999, p. 822). If people had rational expectations, we would therefore expect that richer people have a higher ▶ quality of life and that, as people become wealthier, their ▶ quality of life improves. Is that really the case?

Richer people clearly have access to a wider variety and larger quantities of goods and services. They also have better access to education and health services and live longer (Wilkison, 1996). They are more likely to afford fun and pleasurable activities and to be able to avoid negative events and persons (Diener, Horwitz, & Emmons, 1985). But how does material well-being compare with subjective indicators of quality of life?

In the past several decades, a large number of studies have looked at the relationship between income and ▶ happiness at a particular point in time and place. The general finding has been that people who have more money are happier on average (Andrews & Withey, 1976; Blanchflower & Oswald, 2004; Campbell, Converse, & Rodgers, 1976; Cummins, 2000; Easterlin, 1974, 1995, 2001a; Larson, 1978). Correlation, however, does not imply causation. It may well be that it is not higher income that makes people happier but rather that happier people earn higher incomes. Studies looking at British lottery winners and people receiving an inheritance found that they reported higher mental well-being in the year following the financial windfall (Brickman, Coates, & Janoff-Bulman, 1978; Smith & Razzell, 1975). This suggests that causation may indeed run from income to ▶ happiness.

While richer people may be happier than poorer people, the difference in their ▶ subjective well-being levels is not as large as one might expect. Brickman et al. (1978) find that despite their sudden increase in wealth, lottery winners are not much happier than people struck by traumas, such as paraplegia or blindness. Diener et al. (1985) look at some of the wealthiest individuals in the United States and find their ▶ happiness levels to be barely above those of

individuals with average income. Moreover, the fact that not everyone in their wealthy group is happier than everyone in their nonwealthy group shows that money may help ▶ happiness, but it does not guarantee it. In general, differences in income explain only a small proportion of the differences in ▶ life satisfaction among persons (Frey & Stutzer, 2002). The influence of money can be mediated by other problems. As Chambers (1997) put it, "amassing wealth does not assure well-being and may diminish it" (p. 1728).

The time series association between improvements in ▶ happiness and income growth has been found to be even more dubious. Countries such as Germany, Denmark, or Italy reported small increases in ▶ life satisfaction while experiencing substantial improvements in per capita GDP during the 1970s and 1980s (Diener & Oishi, 2000). Furthermore, in countries such as the United States, the United Kingdom, Belgium, or Japan, which in the past century have experienced sharp increases in GDP per capita, average ▶ happiness has stayed virtually constant or has even declined during the same period (Blanchflower & Oswald, 2004; Diener & Oishi, 2000; Easterlin, 1974, 1995, 2005; Kenny, 1999; Lane, 1998; Myers, 2000). This lack of a positive association between income growth and increases in ▶ happiness persists when individual characteristics are controlled for (Blanchflower & Oswald, 2004) or when looking at the average ▶ happiness of a given cohort as income grows over the life cycle (Easterlin, 2001b). Similar conclusions were reached when looking at the very rich. Lottery winners experience a jolt in their ▶ happiness levels, but the euphoria eventually fades (Argyle, 1986; Brickman et al., 1978). In general, those whose income has increased over the previous decade are not happier than those whose income has not increased (Diener, Sandvik, Seidlitz, & Diener, 1993).

Numerous studies have looked into why increases in income do not automatically lead to commensurate increases in ▶ subjective well-being. One of the psychological processes mediating the relationship between income and happiness is ▶ adaptation. People adapt to higher levels of income, and as a result, the ex post evaluation of a higher income is smaller than its ex ante evaluation. This is why what rich people consider a "sufficient" income is higher than what poor people consider a "sufficient" income (Frey & Stutzer, 2002). In determining their income aspirations, people also compare their incomes not with what they need to live comfortably but with those around them who have the most. Because people look upward when making comparisons (Stutzer, 2004), even the relatively affluent fall victims to the phenomenon of "relative deprivation" (Martin, 1981; Williams, 1975) by comparing themselves with the very rich.

As a result of ▶ adaptation and ▶ social comparison, increases in income tend to be accompanied by increases in aspirations (Diener, 2000). Because a person's level of ▶ happiness has less to do with the objective amount of money they have and more to do with how this amount compares with their aspiration levels and these aspirations in turn increase as income increases, people labor on a "▶ hedonic treadmill," as Brickman et al. (1978) put it.

Whether rich people are happier has also been found to depend on the wealth of a country (Diener & Biswas-Diener, 2002; Veenhoven, 1991). In poor countries, where low income may not be enough to satisfy basic human needs, being relatively well-off does predict greater well-being (Argyle, 1999). Once life's necessities are assured though, more money provides ▶ diminishing returns on ▶ happiness, which is why ▶ affluence matters surprisingly little (Myers, 2000). As Lykken (1999, p. 17) observed "People who go to work in their overalls and on the bus are just as happy, on average, as those in suits who drive to work in their own Mercedes."

## Conclusion

At a point in time, richer people tend to be happier than poorer people, but the difference is not as large as one might expect. Over time, as their income increases, people do not necessarily become happier. Psychological processes of ▶ adaptation and ▶ social comparison can lead to an increase in aspirations which offsets the

positive effect of the increase in income. This is more likely to happen in richer countries, where even people with low incomes are generally able to satisfy their basic necessities.

## Cross-References

- Adaptation
- Adaptation-Level Theory
- Affluence
- Aspiration Theory
- Aspiration-Achievement Gap
- Basic Needs
- Diminishing Returns
- Happiness
- Hedonic Adaptation
- Income Influence on Satisfaction/Happiness
- Life Satisfaction
- Personal Income, Satisfaction with
- Quality of Life (QOL)
- Rational Choice Theory
- Relative Deprivation Theory
- Relative Income and Reference Group Behavior
- Satisfaction with Life as a Whole
- Social Comparison Theory
- Subjective Well-being

## References


Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: America's perception of life quality*. New York: Plenum.

Argyle, M. (1986). *The psychology of happiness*. London: Methuen.

Argyle, M. (1999). Causes and correlates of happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 353–373). New York: Russell Sage.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of Public Economics, 88*, 1359–1386.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology, 36*, 917–927.

Campbell, A. (1981). *The sense of well-being in America*. New York: McGraw-Hill.

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life*. New York: Russell Sage.

Chambers, R. (1997). Editorial: Responsible well-being – A personal agenda for development. *World Development, 25*, 174301754.

Csikszentmihalyi, M. (1999). If we are so rich, why aren't we happy? *American Psychologist, 54*, 821–827.

Cummins, R. A. (2000). Personal income and subjective well-being: A review. *Journal of Happiness Studies, 1*, 133–158.

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist, 55*, 34–43.

Diener, E., & Biswas-Diener, R. (2002). Will money increase subjective well-being? A literature review and guide to needed research. *Social Indicators Research, 57*, 119–169.

Diener, E., Horwitz, J., & Emmons, R. A. (1985). Happiness of the very wealthy. *Social Indicators Research, 16*, 263–274.

Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well-being across nations. In E. Diener & E. M. Suh (Eds.), *Culture and subjective well-being* (pp. 185–218). Cambridge, MA: MIT Press.

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute? *Social Indicators Research, 28*, 195–223.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth* (pp. 89–125). New York: Academic Press.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization, 27*, 35–47.

Easterlin, R. A. (2001a). Income and happiness: Towards a unified theory. *The Economic Journal, 111*, 465–484.

Easterlin, R. A. (2001b). Life cycle welfare: Evidence and conjecture. *Journal of Socio-Economics, 30*, 31–61.

Easterlin, R. A. (2005). Diminishing marginal utility of income? Caveat Emptor. *Social Indicators Research, 70*, 243–355.

Frey, B. S., & Stutzer, A. (2002). What can economists learn from happiness research? *Journal of Economic Literature, 40*, 402–435.

Kenny, C. (1999). Does growth cause happiness, or does happiness cause growth? *Kyklos, 52*, 3–26.

Lane, R. E. (1998). The joyless market economy. In A. Ben-Ner & L. Putterman (Eds.), *Economics, values, and organization* (pp. 461–488). Cambridge, UK: Cambridge University Press.

Larson, R. (1978). Thirty years of research on the subjective well-being of older Americans. *Journal of Gerontology, 33*, 109–125.

Lykken, D. (1999). *Happiness*. New York: Golden Books.

Martin, J. (1981). Relative deprivation: A theory of distributive injustice for an era of shrinking resources. *Research in Organizational Behavior, 3*, 53–107.

Myers, D. G. (2000). The funds, friends, and faith of happy people. *American Psychologist, 55*, 56–67.

Smith, S., & Razzell, P. (1975). *The pools winners*. London: Caliban Books.
Stutzer, A. (2004). The role of income aspirations in individual happiness. *Journal of Economic Behavior and Organization, 54*, 89–109.
Veenhoven, R. (1991). Is happiness relative? *Social Indicators Research, 24*, 1–34.
Wilkison, R. G. (1996). *Unhealthy societies: The afflictions of inequality*. London: Routledge.
Williams, R. M., Jr. (1975). Relative deprivation. In L. A. Coster (Ed.), *The idea of social structure: Papers in Honor of Robert K Merton* (pp. 355–378). New York: Harcourt Brace Jovanovich.

## Right and Wrong

▶ Morality and Well-Being

## Rights and Quality of Life

▶ Human Rights

## Rights of the Child

▶ Child Rights

## Rights-Based Development

▶ Democracy and Development in Mexico and Chile

## Riparian Zone

▶ Wetland(s)

## Risk Factors

▶ At-Risk Children

## Risk of Poverty

▶ Low Income

## Risky Lifestyles and Self-Control

Robert Svensson[1] and Lieven J. R. Pauwels[2]
[1]Malmö University, Malmö, Sweden
[2]Department of Criminal Law and Criminology, Ghent University Malmö University, Gent, Belgium

### Synonyms

Integrated lifestyle exposure theory; Theory of low self-control

### Definition

Risky lifestyles and low self-control have been identified as two major and stable correlates of offending and ▶ victimization, two factors that can seriously underpin the ▶ quality of life of people. Having a risky lifestyle and low levels of self-control are two key mechanisms that may contribute to our understanding of the relationship between offending and victimization. Many studies give empirical evidence for this finding.

#### The Concept of Lifestyle Risk

Lifestyles were originally considered of importance as a causal mechanism in the explanation of victimization, while some more recent studies actually suggest that lifestyles are of much more importance in explaining individual differences in offending. One can argue that much has happened since the classic lifestyle model of victimization. The concept of lifestyles has been fine-tuned, new measures of lifestyles have been developed that go beyond background variables as proxies, and the role of lifestyles has been expanded and incorporated into theories of

offending. Indeed, recent developments within theories of offending increasingly point to the importance of ▶crime inducing settings as situational triggers to explain involvement in offending. In doing so, they very similarly point to the role of lifestyles and daily routines (Wikström & Butterworth, 2006). Previous studies have by and large studied lifestyles in a restricted meaning, referring exclusively to vocational and leisure activities (e.g., Nofziger & Kurtz, 2005; Osgood & Anderson, 2004; Vazsonyi, Pickering, Junger, & Hessing, 2001). Some contemporary scholars use the concept of lifestyle risk in a broader sense and consider lifestyles to be a multidimensional concept, which consists of unstructured routines, peer delinquency, and the use of alcohol (Svensson & Pauwels, 2010; Wikström & Butterworth, 2006).

### The Concept of Low Self-Control

Low self-control is seen as a trait, or a summary construct of individual traits including impulsivity, insensitivity, risk-taking, and shortsightedness, that has "a considerable tendency/. . ./ to come together in the same people, /. . ./persist through life" Gottfredson & Hirschi (1990), pp. 90–91) and which are established very early in life. The theory explains individual differences in crime involvement as variation "in the extent to which [individuals] are vulnerable to the temptations of the moment" (ibid. p. 87). Individuals with low self-control are more vulnerable to temptations of the moment because they fail "to consider the negative or painful consequences of [their] acts and therefore are more likely to engage in acts of crime".

## Description

### The Theory of Lifestyle Risk, Victimization, and Offending

The integrated lifestyle exposure theory (Hindelang, Gottfredson, & Garofalo, 1978) is probably the best known classic theory that explicitly explains individual differences in personal victimization. This theory has been originally developed to explain why the distribution of victimization is unequal and varies by background characteristics. According to the theory, individuals are differentially involved in leisure activities, vocational activities, and professional activities due to their different backgrounds. Such background characteristics that refer to age, race, income, and marital status bring about differences in role expectations (what to expect in life) and constraints (what can be achieved given a certain structural reference). Individuals are restricted by their backgrounds in realizing or making possible achievements. Individuals react and thus adapt differentially to the structural conditions they are part of. Such adaptations bring about observable differences in lifestyles. In this classic model, lifestyle is seen as an indicator of one's daily routines. The original theory focused on vocational-, professional-, and leisure-related activities. Such activities set the stage for exposure to offending in at-risk situations. Differences in lifestyles cause individuals to be differentially exposed to the environmental settings they encounter. For example, adolescents who have unstructured routines and spend a lot of their time on streets, in absence of parental or social control, are more likely to encounter offenders, as also offenders spend a significant part of their leisure time in these settings. Furthermore, differences in lifestyles lead to different associations with (significant) others. For example, adolescents that hang around on the streets are far more likely to get in contact with crime-prone individuals. The classical lifestyle model has some weak and strong points. One strong point is the fact that lifestyles are seen as an indicator of exposure to environmental settings in which victimization is more likely to occur. Lifestyles are thus seen as a situational concept that contributes to our understanding of victimization. One weak point is that the model is rather vague in how lifestyles are developed and how individual differences in lifestyles are brought about. The classical model overemphasized structural and cultural differences, thereby ignoring the role of individual characteristics, such as an individual's level of propensity to offend. A risky lifestyle has

not only been related to victimization but also to offending behavior. The idea that individuals mostly interact with people that are rather similar to each other is the most typical explanation for the relationship between offending and victimization: the risk of victimization is directly in proportion with the amount of the characteristics they share with perpetrators (Hindelang et al.). Thus, in that case, controlling for common characteristics, such as common demographic background variables, deviant lifestyles, and propensity to offend, would make the zero-order correlation between offending and victimization spurious.

### The Theory of Low Self-Control in a Nutshell

The question why some people restrain from committing acts of crime is not only a question raised by criminologists. Many sociologists consider the same question from another point of view: they regard the process of socializing as a continuing key issue in sociology. As originally conceptualized by Gottfredson and Hirschi (1990, p. 232), low self-control is argued to be "the individual-level cause of crime," i.e., low self-control is theorized to be the primary explanation for criminal behavior. The effects of all other theoretical constructs to explain offending are spurious when self-control is entered into the equation. While several scholars have issued critiques of the theory (Akers, 1991; Barlow, 1991; Benson & Moore, 1992; Cohen & Vila, 1996; Geis, 2000; Reed & Yeager, 1996; Tittle, 1991), it also attained considerable empirical support. Low self-control has consistently been found to be a modest to strong correlate of both crime and analogous acts of ▶ deviant behavior, including substance use and/or abuse. Ineffective child rearing (inconsequent parenting) seems to be the major cause of low self-control according to the authors. Self-control is established early in life as a consequence of the socialization process. Gottfredson and Hirschi stress the importance of parental supervision, monitoring, discipline, and affection. In a previous study on the importance of family processes in the development of self-control, Hope, Grasmick, and Pointon (2003) and Vazsonyi and Belliston (2007) noted that only few studies have examined the causes of low self-control consistent with self-control theory (see also Pratt et al., 2004). In other words, many studies to date have focused on the self-control-offending relationship and have not included aspects of the ▶ family structure and social bonds in empirical tests of the theory. This is so despite the fact that Gottfredson and Hirschi clearly point to the importance of the family as the primary socializing agent responsible for the development of self-control. The family is the primary unit in which children learn ▶ values and ▶ attitudes that guide them throughout their lives. In a cross-cultural and cross-national test of self-control theory, Vazsonyi and Belliston were able to demonstrate the relationships between the family processes of monitoring, closeness and support, self-control, and offending. Vazsonyi and Belliston demonstrated both direct and indirect effects of socialization variables on offending. Self-control was a key mediator of the relationship between family social processes and offending, while still direct effects of family mechanisms on offending were found. Families do not only differ in levels of child rearing but also differ in their family structure. Some children grow up in one-parent families, while others grow up in families with two caretakers. Growing up in a one-parent family may be an indicator of experienced family disruption and may therefore have a lasting impact on the ways these children grow up and develop self-control. Parents act as agents of informal control, and when a breakdown in the family structure occurs, it may be a consequence that these children get involved in antisocial behavior.

### Lifestyle Risk, Low Self-Control, and the Relationship Between Offending and Victimization

The idea that individuals mostly interact with people that are rather similar to each other is the most typical explanation for the relationship between offending and victimization: the risk of victimization is directly in proportion with the amount of the characteristics they share with perpetrators (Hindelang et al., 1978). Thus, in

that case, controlling for common characteristics, such as common demographic background variables, deviant lifestyles, and propensity to offend, would make the zero-order correlation between offending and victimization spurious. Different theoretical models have been suggested for the explanation of the association between victimization and offending. As stated above, the relation between victimization and offending can primarily be explained only by the fact that offenders and victims have a common structural background. This explanation supports on the assumption that some daily life patterns do increase the chance to offend as well as the chance of becoming a victim, which is called by Hindelang et al. the "principle of homogamy." The link between offending and victimization is in that case a spurious relationship (Wittebrood & Nieuwbeerta, 1999). A second explanation is offered by self-control theory that argues that low self-control is a common cause of offending and victimization. For example, people that have low self-control are more frequently involved in accidents but are also more commonly involved in offending. Self-control theory in its purest form would argue that the relationship between offending and victimization is spurious. A third explanation is that offending actively contributes to the chance of becoming a victim. In that case, offending is seen as one causal mechanism in explaining victimization "because of the motives, vulnerability, or culpability of people involved in these activities" (Jensen & Brownfield, 1986). Delinquent lifestyles of offenders place them more in rough environments and risky situations, and the interactions with other (potential) offenders may in turn increase offenders' chance of becoming a victim (Bjarnason, Sigurdardottir, & Thorlindsson, 1999; Shaffer & Ruback, 2002; Wittebrood & Nieuwbeerta, 1999). Smith et al. (2001) argue that incidents of offending and victimization are "the outcome of an interactive process, in which both the eventual victim and the eventual delinquent play a part" (Smith et al. 2001: 76). The idea that lifestyles are causally related to victimization can be found in sayings such as "he who lives by the sword shall die by the sword." Most scholars seem to consider offending as one causal mechanism that explains victimization, while the relationship between both characteristics can be non-recursive, i.e., there may exist feedback loops between offending and victimization. Developmental processes may lead from victimization to offending and back in different ways. Fourth, the association between victimization and offending can also be explained by the fact that repeated trauma of violence-exposed children can lead to processes of ▶ anger, so they may participate in high-risk behaviors or associate with offenders – those who place others at risk (Bailey & Whittle, 2004). Offenders may become the target of revenge to compensate for previous suffering, or victimization can lead to offending if the victim has the feeling of being a victim of an injustice, which may motivate the victim to offend themselves in turn.

## Discussion

The results of many studies, including our own studies (Pauwels & Svensson, 2011), clearly challenge theories of victimization that depart from the idea that lifestyle risk and low self-control are key determinants of victimization. The strong and independent effect of lifestyle risk is highly similar to the findings reported in a key study on lifestyles and criminal behavior (Wikström & Butterworth, 2006) and makes a strong plea for the further development of situational models of offending that take a closer look at the role of lifestyles. There is some support for the independent effect of victimization on offending, but the role of propensity to offend and lifestyle risk on offending are stronger. Earlier we concluded from regression models that the standardized effect of offending on victimization is stronger than the effect of victimization on offending in both settings. Although results from cross-sectional statistical analyses never allow for causal interpretation, such results are highly suggestive for the idea that the direct effect of offending on victimization is stronger than the direct effect of victimization of offending. Future studies on the relation between offending and victimization should take the following aspects

into account. Theories that explain individual differences in victimization have been criticized by several authors on different topics. For example, some kinds of victimization that occur in private space like incest and domestic violence cannot be explained by these theories (Finkelhor, 2007). Besides that, these theories analyze only partially the role of human action and structural constraints in victimization (Spalek, 2006; Walklate, 2007b). There still is a lacuna in our knowledge of understanding what other mechanisms besides low self-control and lifestyle risk may increase the risk of victimization. The integrated model of low self-control and lifestyle risk fits offending better than it fits victimization. So this suggests once again that there exist other mechanisms that explain the risk of victimization. In short, especially the explanation of victimization needs to move beyond theories of lifestyle and low self-control (Singer, 1981; Finkelhor & Asdigian, 1996; Wittebrood & Nieuwbeerta, 1999). Based on our own studies among adolescents, we argue that we need to rethink the etiology of victimization and find out what other mechanisms that can explain the correlation between offending and victimization of adolescents. Scholars should avoid to fall back into what is called by Van Dijk (2008) "the guilty conscience of victimology" by blaming the victim or awaking the impression some victims provoke their victimization. Thus, it is ambiguous and even dangerous to generalize all victimization among adolescents as a consequence of offending, lifestyle risk, and low self-control: this does not recognize the generally involuntary and high-risk nature of adolescent environment.

## Cross-References

► Quality of Life Research

## References

Akers, R. (1991). Self-control as a general theory of crime. *Journal of Quantitative Criminology*, *7*, 201–211.

Bailey, S., & Whittle, N. (2004). Young people: Victims of violence. *Current Opinion in Psychiatry*, *17*, 263–268.

Barlow, H. (1991). Explaining crime and analogous acts, or the unrestrained will grab at pleasure wherever they can. *Journal of Criminal Law and Criminology*, *82*, 229–242.

Benson, M. L., & Moore, E. (1992). Are white-collar and common offenders the same? *Journal of Research in Crime and Delinquency*, *29*, 251–272.

Bjarnason, T., Sigurdardottir, T. J., & Thorlindsson, T. (1999). Human agency, capable. Guardians, and structural constraints: A lifestyle approach to the study of violent victimization. *Journal of Youth and Adolescence, 28*, 105–119.

Cohen, L., & Vila, B. (1996). Self-control and social control: An exposition of the Gottfredson-Hirschi/Sampson-Laub debate. *Studies in Crime and Crime Prevention*, *5*, 125–150.

Finkelhor, D., & Asdigian, N. L. (1996). Risk factors for youth victimization: Beyond a lifestyles/routine activities theory approach. *Violence and Victims*, *11*(1), 3–19.

Finkelhor, D. (2007). Developmental victimology: The comprehensive study of childhood victimizations. In R. C. Davis, A. J. Lurigio, & S. Herman (Eds.), *Victims of crime* (pp. 9–34). Los Angeles: Sage Publications.

Geis, G. (2000). On the absence of self-control as the basis for a general theory of crime: A critique. *Theoretical Criminology*, *4*, 35–53.

Gottfredson, M., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford California Press.

Hindelang, H. J., Gottfredson, M. R., & Garofalo, S. (1978). *Victims of personal crime: An emperical foundation for a theory of personal victimization*. Cambridge, UK: Ballinger Publishing Company.

Hope, T. L., Grasmick, H. G., & Pointon, L. J. (2003). The family in Gottfredson and Hirschi's general theory of crime: Structure, parenting and self-control. *Sociological Focus*, *36*, 291–311.

Jensen, G. F., & Brownfield, D. (1986). Gender, lifestyles, and victimization: Beyond routine activity. *Violence and Victims, 1*(2), 85–99.

Nofziger, S., & Kurtz, D. (2005). Violent lives: A life style model linking exposure to violence to juvenile violent offending. *Journal of Research in Crime and Delinquency, 42*, 3–26.

Osgood, D. W., & Anderson, A. L. (2004). Unstructured socializing and rates of delinquency. *Criminology, 42*, 519–549.

Pauwels, L., & Svensson, R. (2011). Exploring the relationship between offending and victimization: What is the role of risky lifestyles and low self-control? A test in two urban samples. *European Journal of Criminal Policy & Research, 17*(3), 163–177.

Pratt, T. C., Turner, M. G., & Piquero, A. (2004). Parental socialization and community context: A longitudinal analysis of the structural sources of low self-control. *Journal of Research in Crime and Delinquency*, *41*, 219–243.

Reed, G., & Yeager, P. (1996). Organizational offending and neoclassical criminology: Challenging the reach of a general theory of crime. *Criminology*, *34*, 357–382.

Shaffer, J. N., & Ruback, R. B. (2002). *Violent victimization as a risk factor for violent offending among juveniles*, *juvenile justice bulletin*. Accessed Jan 08, 2012, from http://www.ncjrs.gov/html/ojjdp/jjbul2002_12_1/contents.html.
Singer, S. I. (1981). Homogeneous victim-offender populations: A review and research implications. *The Journal of Criminal Law & Criminology*, *72*, 779–788.
Smith, D. J., McVie, S., Woodward, R., Shute, J., Flint, J., & McAra, L. (2001). *The Edinburgh Study of Youth Transitions and Crime*: *Key findings at ages 12 and 13* (209 p). Retrieved March 18, 2009, from http://www.law.ed.ac.uk/cls/esytc/findreport/wholereport.pdf
Spalek, B. (2006). *Crime victims*: *Theory*, *policy and practice*. Houndmills, UK: Palgrave Macmillan.
Svensson, R., & Pauwels, L. (2010). Is a risky lifestyle always "risky"? The interaction between individual propensity and lifestyle risk in adolescent offending: A test in two urban samples. *Crime & Delinquency*, *56*, 608–626.
Tittle, C. (1991). A general theory of crime: A book review. *American Journal of Sociology*, *96*, 1609–1611.
van Dijk, J. J. M. (2008). *Slachtoffers als zondebokken*. Apeldoorn/Antwerpen, Belgium: Maklu.
Vazsonyi, A., & Belliston, L. (2007). The family -> low self-control -> deviance. A cross-cultural and cross-national test of self-control theory. *Criminal Justice and Behavior, 34*(4), 505–530.
Vazsonyi, A., Pickering, L., Junger, M., & Hessing, D. (2001). An empirical test of a general theory of crime: A four nation comparative study of self-control and prediction of deviance. *Journal of Research in Crime and Delinquency, 38*, 91–131.
Walklate, S. (2007b). *Understanding criminology*: *Current theoretical debates*. Maidenhead, UK: Open University Press.
Wikström, P.-O. H., & Butterworth, D. A. (2006). *Adolescent crime: Individuals differences and lifestyles*. Devon, England: Willan Publishing.
Wittebrood, K., & Nieuwbeerta, P. (1999). Wages of sin? The link between offending, lifestyle and violent victimization. *European Journal on Criminal Policy and Research, 7*, 63–80.

## Rivalry, Sibling

▶ Violence in USA, Sibling

## River Basins

▶ Watershed(s)

## RMQoL (Radex Model of Quality of Life)

▶ Quality of Life, Radex Model

## Road Noise

▶ Traffic Noise Abatement

## Robust

▶ Robust Statistical Tests

## Robust Statistical Tests

Donald W. Zimmerman
Carleton University, Surrey, BC, Canada

### Synonyms

Analysis of residuals; Analysis of variance; Linear regression model; Nonparametric statistics; Outliers; Robust

### Definition

Statistical tests are widely used in social, behavioral, and health science research to test hypotheses. It is well known that these statistical tests depend on assumptions such as normality, equal variances, and independence of the observations. This entry focuses on whether widely used statistical tests lead to correct statistical decisions, even when assumptions made in the mathematical derivation of the tests are not strictly fulfilled.

## Description

This entry examines some unexpected effects that occur when two or more assumptions underlying commonly used parametric significance tests are violated at the same time. It also emphasizes that nonparametric methods that are robust under non-normality often are ineffective under heterogeneity of variance. Furthermore, violation of independence, not frequently studied, has pronounced effects on all commonly used significance tests, parametric and nonparametric. The problems that arise are illustrated by simulations, using the Student *t* test and the Wilcoxon-Mann–Whitney rank-sum test applied to data from normal and nine non-normal distributions, under conditions where homogeneity of variance and independence of sample observations also are violated.

Extensive simulation studies over many years have addressed the *robustness* of statistical significance tests. The term refers to the extent to which a significance test leads to correct statistical decisions, even when assumptions made in the mathematical derivation of the test are not strictly fulfilled. It is important to recognize that the *power* of the test, or freedom from type II errors, is as important as type I error rates. Simulation studies have revealed that a test sometimes can perform well in holding type I error rates close to the nominal significance level but be seriously lacking in power to detect alternatives. Furthermore, *inflation* of the probability of rejecting $H_0$, that is, spuriously elevated *power*, can be as damaging as insufficient power.

Significance tests of many types make various assumptions depending on their specific purpose and the experimental designs to which they are applied, but many widely used tests are based on three assumptions which are sometimes questionable in practical research. These are (1) *normality*, (2) *homogeneity of variance*, and (3) *independence of sample observations*. For example, the Student *t* test, analysis of variance (ANOVA), and related tests of significance of correlation and regression all assume normality. When two or more samples are compared, the same tests assume *homogeneity of variance*, or equality of variances of treatment groups. And almost all significance tests assume *independence* of observations, a familiar concept in probability and statistics that enters into the definition of a random sample.

In the early part of the last century, it was often assumed that many of the commonly used parametric tests based on normal theory are *robust*, that is, that they usually lead to the correct statistical decisions despite violation of assumptions that are not too severe (see, e.g., Boneau, 1960). Over the years, however, a wide variety of studies, both Monte Carlo studies and mathematical derivations, disclosed that the type I error rates and power of the *t* test, the ANOVA *F* test, and others could be extensively compromised when normality and the other basic assumptions are not fulfilled (see, e.g., Blair & Higgins, 1980, 1985, Hodges & Lehmann, 1956; Hsu, 1938; Randles & Wolfe, 1979).

### Nonindependence of Separate Violations of Assumptions

The introduction of nonparametric statistics was found to protect against failure of the normality assumption under many conditions. However, nonparametric tests also make their own special assumptions that are not always realistic in practice, and the tests sometimes perform poorly despite their success in overcoming non-normality. For example, some widely used nonparametric tests based on ranks, including the Wilcoxon-Mann–Whitney test, the Wilcoxon signed-ranks test, and the Kruskal-Wallis test, which often are used in place of the *t* and *F* tests, are not at all effective against variance heterogeneity. A test that does a good job in protecting against violation of one assumption does not necessarily protect against violation of another.

There exists a vast literature of studies that have explored the type I and type II errors of tests under violation of assumptions, but it is not always possible for researchers to obtain a clear picture of what can be expected in practice. One problem is that simulation studies have tended to focus on one problem at a time, that is, examining what happens to ANOVA, when normality is violated, without considering effects

of heterogeneity of variance, or vice versa. Similarly, modifications of the *t* and *F* tests have been found to be effective for normal distributions when variances are unequal, but do not perform well under non-normality (Alexander & Govern, 1994; Cressie & Whitford, 1986; Fagerland & Sandvik, 2009; Keselman, Wilcox, Taylor, & Kowalchuk, 2000; Satterthwaite, 1946; Scheffé, 1970; Stonehouse & Forrester, 1998; Welch, 1938, 1947).

The violation that has been least investigated, *nonindependence of sample observations*, can bring about substantial changes in type I and type II error probabilities that exceed those produced by non-normality and variance heterogeneity. Relatively little has been done in this area, and nonparametric methods based on ranks do not effectively handle the problem. Modifications of the *t* and *F* tests based on formulas that explicitly contain correlation terms have not been widely explored (see, however, Zimmerman, Williams, & Zumbo, 1993; Zumbo & Zimmerman, 1991).

Furthermore, it is important to recognize that there is interaction with respect to violation of assumptions. That is, heterogeneity of variance can produce a certain result under normality but an entirely different result for various non-normal distributions. Or one non-normal distribution can lead to results quite different from another, depending on skewness and kurtosis, the presence of outliers, and so on. Furthermore, robustness also depends on parameters such as sample sizes, the degree of difference between means, and the significance level.

## Simulation Procedures

This entry attempts to provide an overall picture of what can happen when the three violations listed above occur both individually and in combination. We examine type II errors and power, in addition to type I error rates, under a variety of non-normal distribution shapes and various degrees of heteroscedasticity, using two familiar two-sample tests, the Student *t* test for independent samples, and the Wilcoxon-Mann–Whitney rank-sum test (WMW). We also look at performance under violation of the independence assumption. It is hoped that these comparisons will provide a more global picture of some changes that are likely to be encountered in practical research studies, where, as Kruskal (1988) remarked, "nature is not necessarily limited to one nuisance at a time."

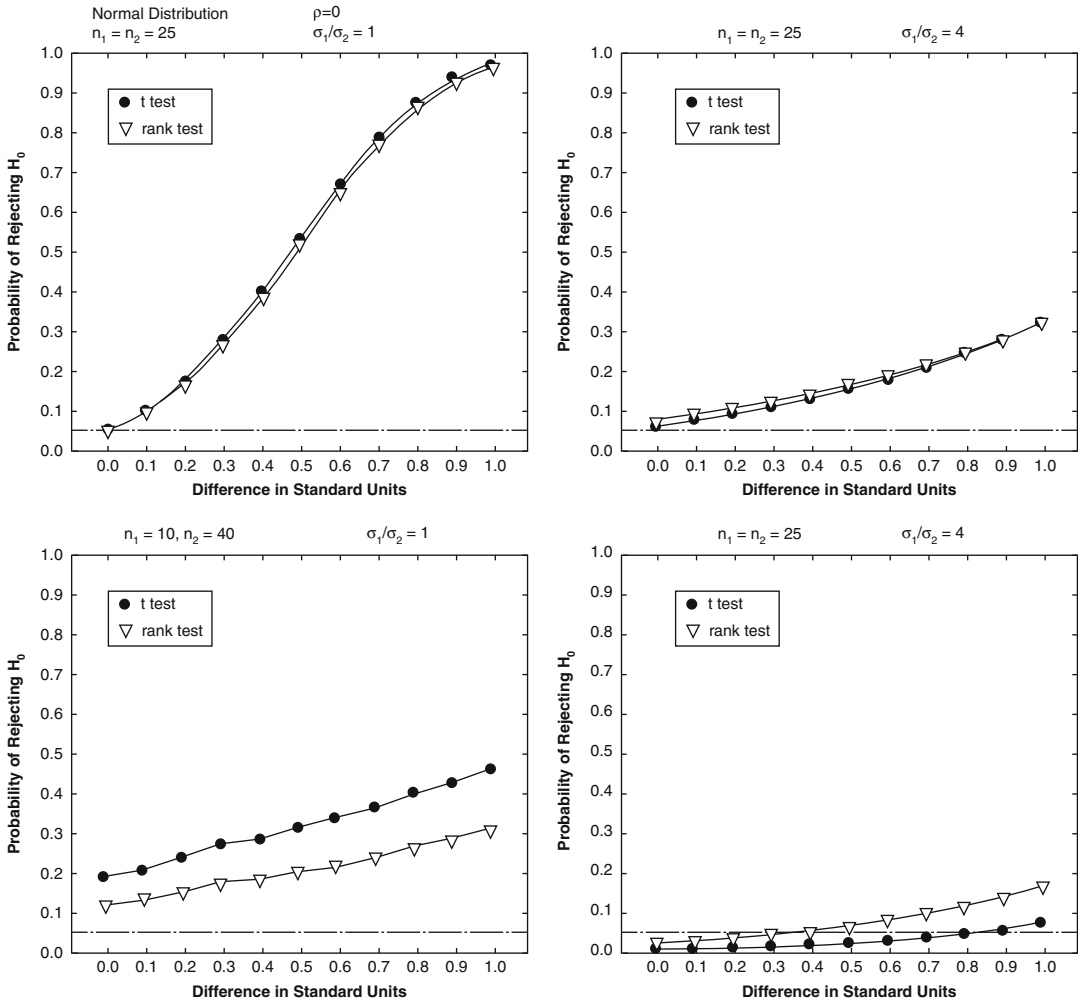
The simulations in this study were programmed using the *Mathematica*, version 4.1, programming language (Wolfram, 1999). Samples from various normal and non-normal distributions were obtained from *Mathematica* statistical add-on packages. Correlations between observations were induced by a method described by Zimmerman et al. (1993). All statistical tests were nondirectional and were performed at the .05 significance level.

## Simulations Based on Two or More Violations of Assumptions

Figures 1 and 2 plot power functions of the independent groups *t* test and the WMW test performed on samples from a normal and an exponential distribution under four conditions: (1) sample sizes equal, $n_1 = n_2$, and standard deviations equal, $s_1 = s_2$; (2) $n_1 = n_2$ and $s_1/s_2 = 4$; (3) $n_1/n_2 = 1/4$ and $s_1/s_2 = 4$ (the smaller standard deviation associated with the larger sample size); and (4) $n_1/n_2 = 4$ and $s_1/s_2 = 4$ (the larger standard deviation associated with the larger sample size).

Figures 1 and 2 show the outcome of these four conditions when sample observations were independent. Figures 3 and 4 are based on the same combinations of ratios of sample sizes and standard deviations, where sample observations are mutually correlated with $\rho = .20$. The 16 graphs in Figs. 1, 2, 3, and 4 therefore represent combinations of the three violations of assumptions, individually, in pairs, or all together. In all cases the difference between the means of the two groups, $m_1 - m_2$, ranged from 0 to 1 s in increments of.1 s.

For the normal and exponential distributions, where sample sizes were equal and $\rho = 0$ (upper left-hand graph in Figs. 1 and 2), the results were essentially the same as found in many simulation studies during the last 30 years. The *t* test slightly dominated the WMW test for the normal distribution, whereas the WMW test dominated the *t*

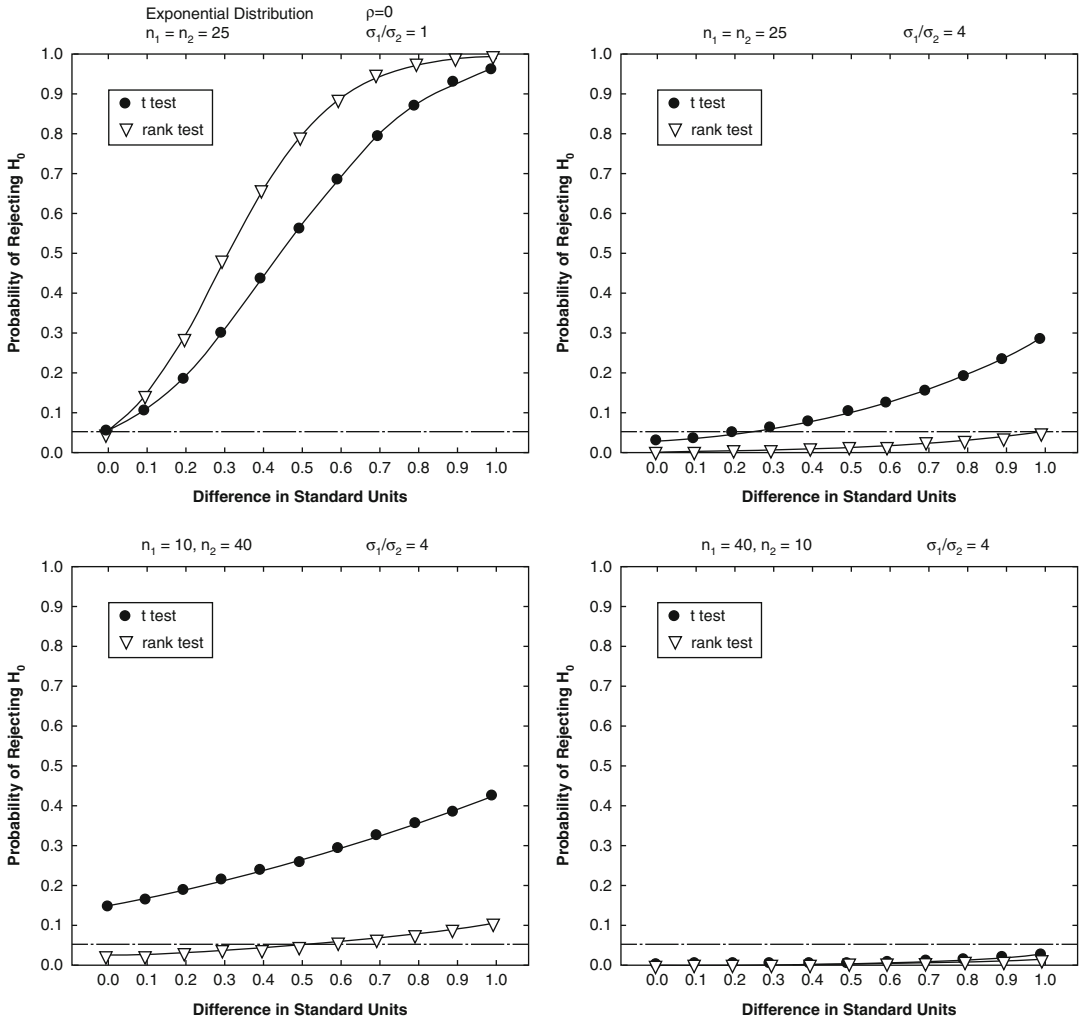



**Robust Statistical Tests, Fig. 1** Power functions of the *t* test and the Wilcoxon-Mann–Whitney test, for a normal distribution, for different ratios of standard deviations of treatment groups and different ratios of sample sizes. Successive sample observations are independent

test to a far greater extent for the exponential distribution. When sample sizes were equal but standard deviations were unequal (upper right-hand graph in Figs. 1 and 2), the power functions of both tests were greatly depressed, and the dominance relation was reversed. That effect has not been widely reported in the literature.

Note that, in the case of the normal distribution, the type I error rates were not noticeably changed by variance heterogeneity, whereas the power of both tests was severely compromised to an extent that neither test would be satisfactory for detecting differences. It has been assumed for a long time that variance heterogeneity has minimal influence when sample sizes are equal, and that is true for type I error rates, but these graphs make it clear that there is also a pronounced effect on type II errors and power.

When both sample sizes and standard deviations were unequal, the type I error rates of the *t* test and the WMW test again were the same as found in many previous studies. If the smaller sample size was associated with the larger standard deviation, the type I error rate of the *t* test was substantially increased above the nominal significance level, and that of the WMW test

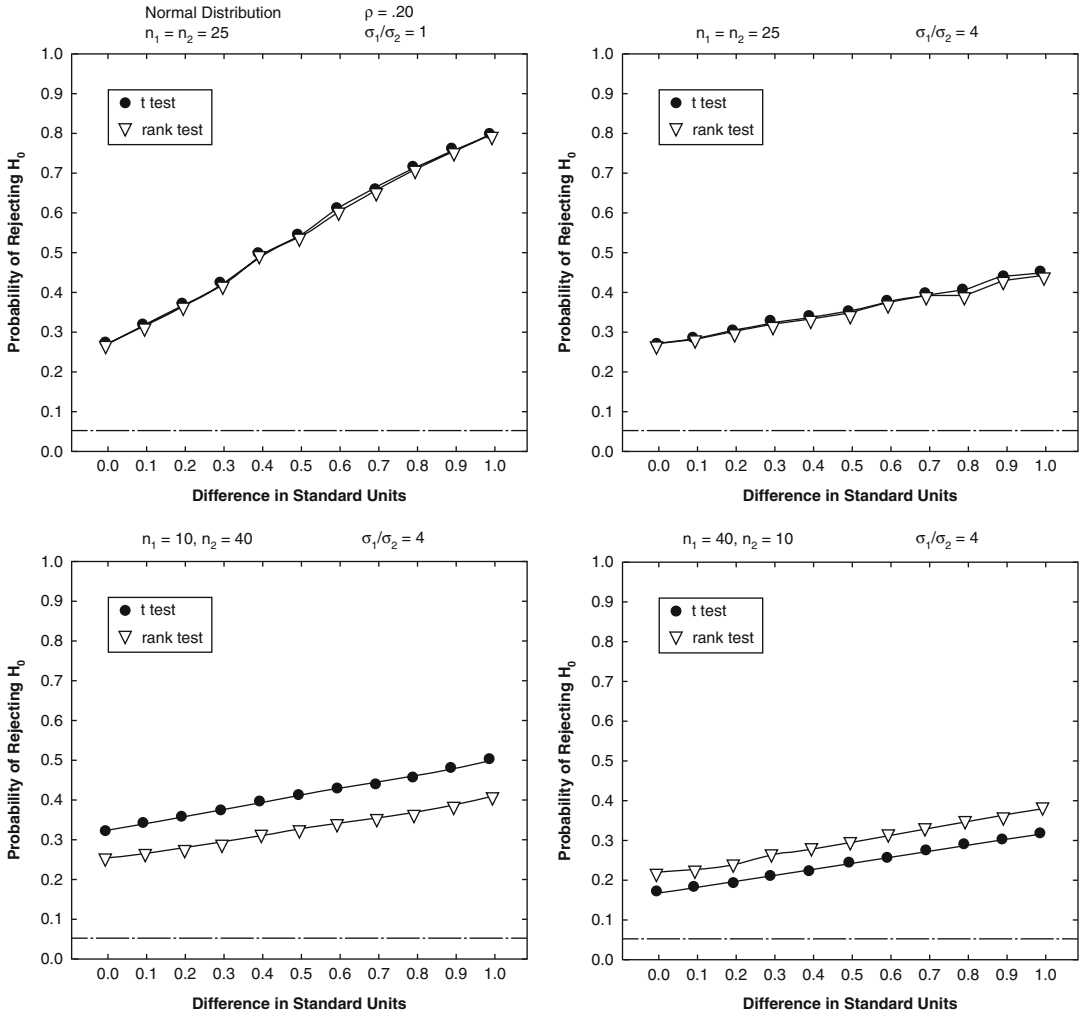



**Robust Statistical Tests, Fig. 2** Power functions of the *t* test and the Wilcoxon-Mann–Whitney test, for an exponential distribution, for different ratios of standard deviations of treatment groups and different ratios of sample sizes. Successive sample observations are independent

was also elevated but to a somewhat lesser extent. On the other hand, the probability of rejecting $H_0$ of both tests remained relatively small over the entire range of values of $m_1 - m_2$. When the larger sample size was associated with the larger standard deviation, the type I error rates of both tests declined below the nominal significance level, that of the *t* test to a greater extent. Again, the entire power functions were affected, and neither test exhibited satisfactory power.

In the case of the exponential distribution, the changes in type I error rates under variance heterogeneity with sample sizes unequal were somewhat less, but the changes in power were magnified, especially for the WMW test. The *t* test had only slight power, and the WMW test virtually no power at all to detect alternatives in all three cases where standard deviations were unequal.

Note that the relations among the functions for the *t* and WMW tests in Figs. 3 and 4 correspond closely to the same relations in Figs. 1 and 2. However, in the former case the curves are considerably elevated. Since the type I error rates, as

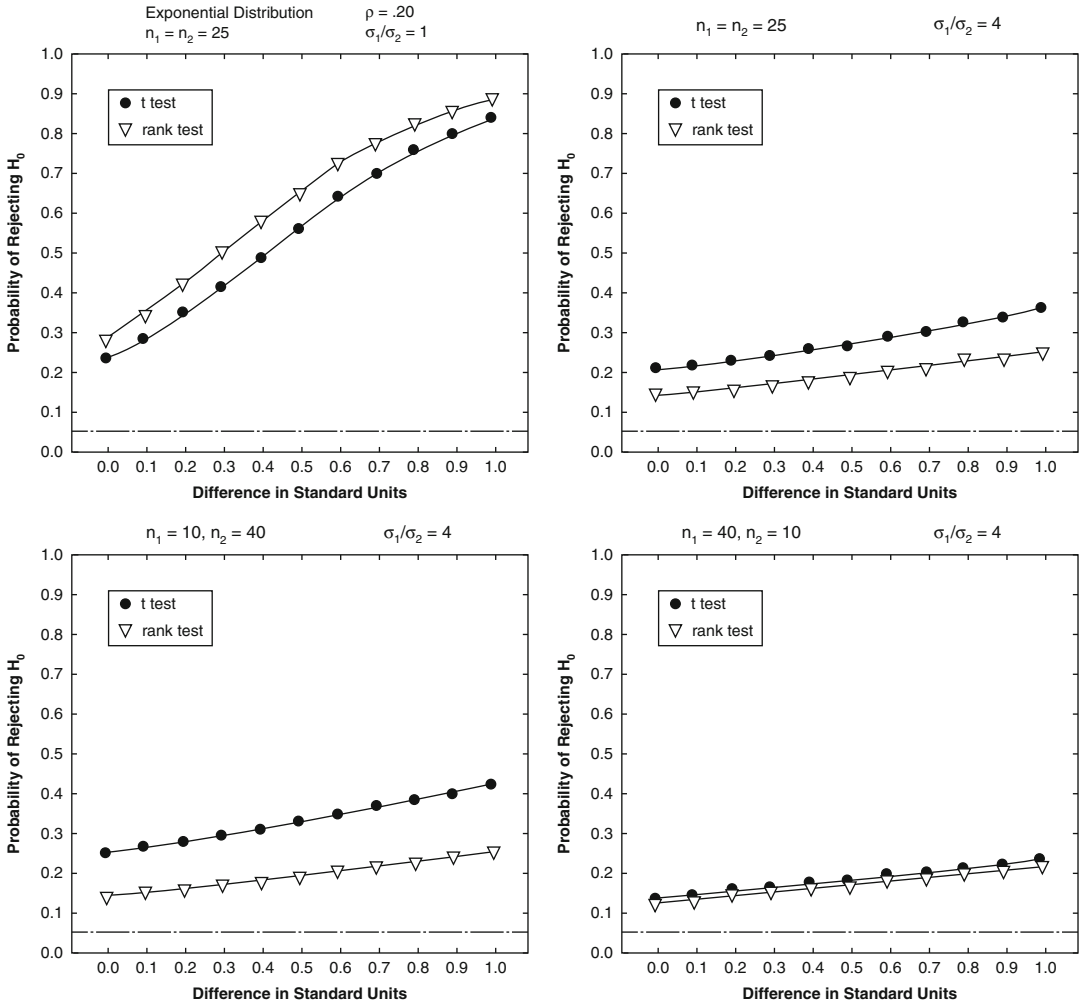



**Robust Statistical Tests, Fig. 3** Power functions of the *t* test and the Wilcoxon-Mann–Whitney test, for a normal distribution, for different ratios of standard deviations of treatment groups and different ratios of sample sizes. Successive sample observations are correlated (ρ = .20)

R

well as power, were increased, these changes in power are of course spurious; that is, rejecting the null hypothesis does not necessarily indicate detection of a true difference between means. Furthermore, the functions for both tests are relatively flat, so that the probability of rejection of $H_0$ for a large difference is not substantially greater than it is for a small or zero difference. Obviously under those conditions the significance tests are unsatisfactory for practical use. When the functions were already inflated because a smaller sample size was associated with a larger variance (lower left-hand section in the figures), the functions were inflated still more under the nonindependence condition. When the *t* test dominated, the same remained true under nonindependence and likewise when the WMW test dominated.

Figures 5, 6, and 7 provide excellent illustrations of the fact that power functions, in addition to type I error rates, are important in assessing the robustness of a test. Figure 5 plots the probability of rejecting $H_0$ as a function of the difference between means, for normal distributions, with the ratio of standard deviations as a parameter. In these cases the smaller standard deviation was

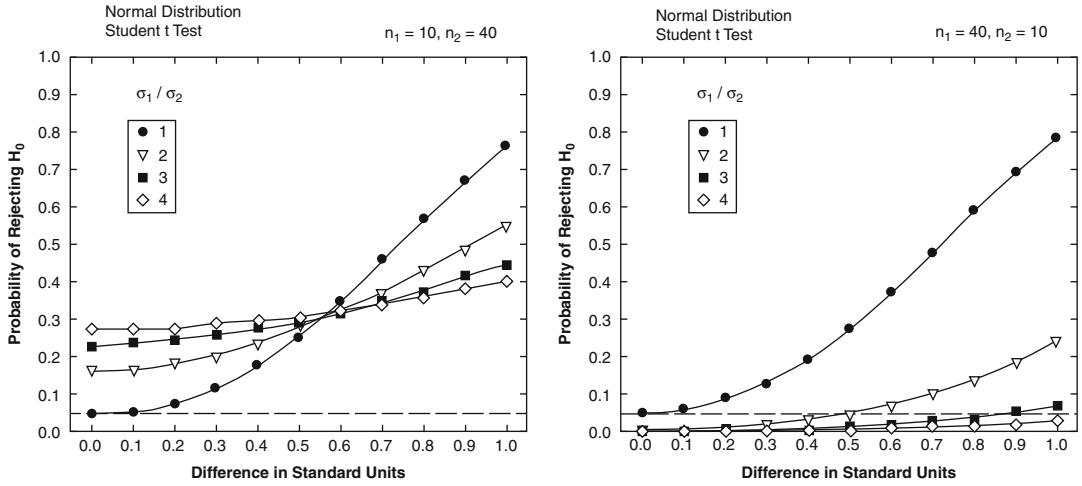



**Robust Statistical Tests, Fig. 4** Power functions of the *t* test and the Wilcoxon-Mann–Whitney test, for an exponential distribution, for different ratios of standard deviations of treatment groups and different ratios of sample sizes. Successive sample observations are correlated ($\rho = .20$)

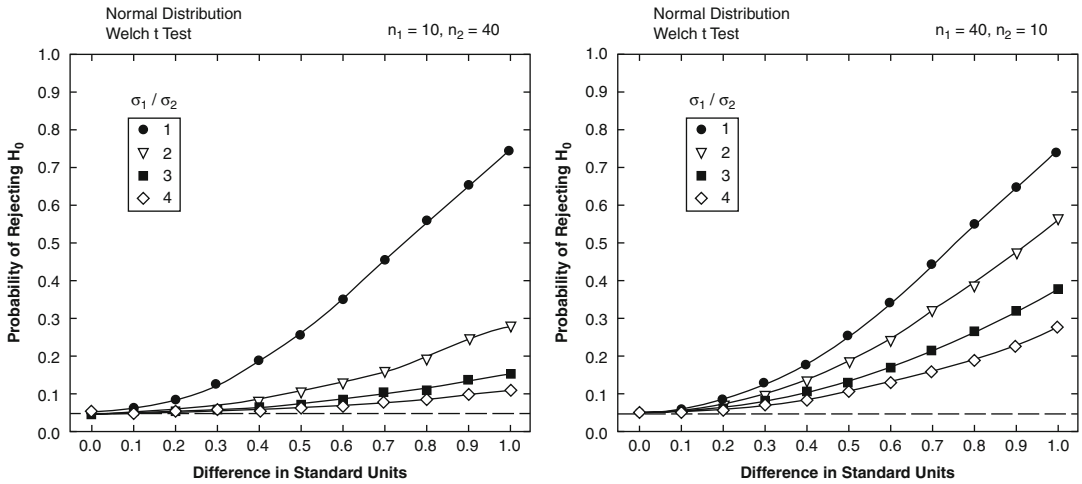
associated with the larger sample size. For all four ratios of standard deviations, the type I error rate was substantially elevated above the.05 significance level, the discrepancy becoming greater as $s_1/s_2$ increased.

Note the *crossover* effect when the difference $m_1 - m_2$ was about.5 s. For the larger differences, the power values were reversed from their previous order. The same crossover effect occurs for the exponential, lognormal, and extreme-value distributions, except that it comes closer to the zero difference on the horizontal axis. The lower section of Fig. 5 shows a similar outcome when $n_1 = 40$ and $n_2 = 10$. In this case the type I error probabilities were depressed below the .05 significance level, and the power functions were drastically modified as before.

Some modifications of the *t* and *F* tests have been used successfully to restore type I error rates for normal distributions under variance heterogeneity (Satterthwaite, 1946; Welch, 1938, 1947). Those methods are not as effective in restoring type I error rates for non-normal distributions, and Fig. 6 reveals that the method breaks down completely for normal distributions when type II errors and power are taken into consideration.




**Robust Statistical Tests, Fig. 5** The probability of rejecting $H_0$ by the $t$ test for a normal distribution as a function of $m_1 - m_2$, for four ratios of standard deviations, for $n_1 = 10$ and $n_2 = 40$ (*upper* section) and $n_1 = 40$ and $n_2 = 10$ (*lower* section)




**Robust Statistical Tests, Fig. 6** The probability of rejecting $H_0$ by the Welch-Satterthwaite modification of the $t$ test as a function of $m_1 - m_2$, for $n_1 = 10$ and $n_2 = 40$ (*upper* section) and $n_1 = 40$ and $n_2 = 10$ (*lower* section)

The upper section of Fig. 6 shows the probability of rejecting $H_0$ by the Welch-Satterthwaite modification of the $t$ test, with $s_1/s_2$ as a parameter, when $n_1 = 10$ and $n_2 = 40$, as in the upper section of Fig. 5. Obviously there is little or no improvement in power; although consistent with known results, the type I error probabilities were restored and the *crossover* effect was eliminated. The lower section of Fig. 6 shows a similar outcome for the case where $n_1 = 40$ and $n_2 = 10$. In that case the depressed type I error probabilities were restored by the modified $t$ test, again consistent with known results, but there was no improvement in the power curves.

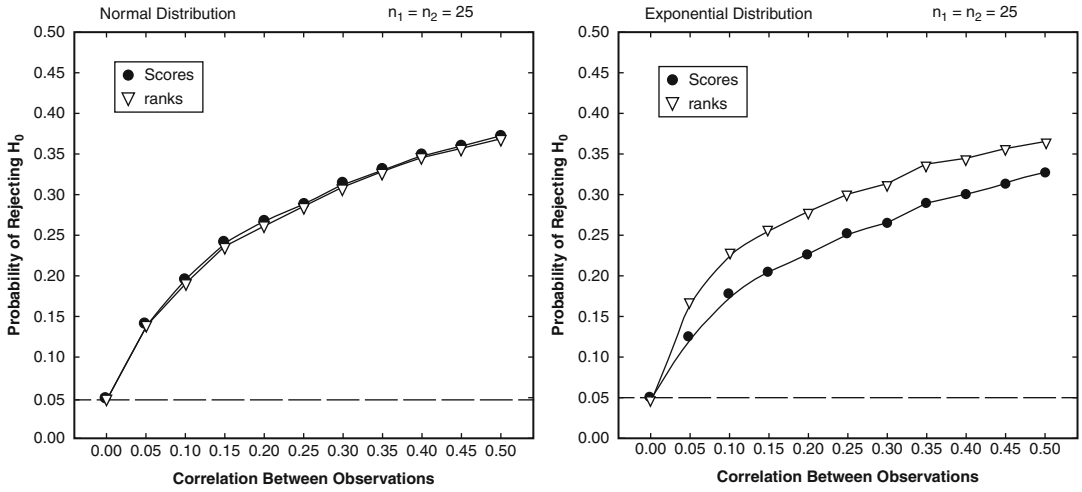
## Additional Non-normal Distributions

Table 1 presents combinations of conditions similar to those in Figs. 1, 2, 3, and 4, for 8 more

**Robust Statistical Tests, Table 1** Probability of rejecting $H_0$ for $m_1 - m_2$ of 0, .5 s, and 1 s and for 8 non-normal distributions, when ratios of standard deviations are 1 and 4, when ratios of sample sizes are 1 and 4, and when observations are independent or nonindependent ($\rho = 0$ and $\rho = .20$) $r = 0$ (independent observations) $r = .20$ (correlated observations)

| | | $n_1 = n_2 = 25$ | | $n_1 = n_2 = 25$ | | $n_1 = 10, n_2 = 40$ | | $n_1 = 40, n_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\sigma_1/\sigma_2 = 1$ | | $\sigma_1/\sigma_2 = 4$ | | $\sigma_1/\sigma_2 = 4$ | | $\sigma_1/\sigma_2 = 4$ | |
| Distribution | $\mu_1 - \mu_2$ | Scores | Ranks | Scores | Ranks | Scores | Ranks | Scores | Ranks |
| | $r = 0$ (independent observations) | | | | | | | | |
| Laplace | 0 | .048 | .046 | .048 | .056 | .188 | .108 | .001 | .015 |
| | .5 | .544 | .650 | .158 | .224 | .322 | .232 | .013 | .098 |
| | 1.0 | .957 | .984 | .347 | .489 | .474 | .402 | .068 | .321 |
| Lognormal | 0 | .045 | .049 | .118 | .008 | .068 | .048 | .000 | .000 |
| | .5 | .639 | .957 | .235 | .025 | .468 | .923 | .001 | .001 |
| | 1.0 | .937 | 1.000 | .434 | .085 | .870 | 1.000 | .018 | .015 |
| Uniform | 0 | .045 | .042 | .053 | .071 | .187 | .128 | .002 | .012 |
| | .5 | .530 | .495 | .145 | .139 | .295 | .177 | .012 | .034 |
| | 1.0 | .867 | .938 | .318 | .242 | .437 | .250 | .057 | .092 |
| Extreme Value | 0 | .048 | .044 | .033 | .015 | .157 | .048 | .011 | .000 |
| | .5 | .546 | .601 | .116 | .052 | .276 | .098 | .004 | .008 |
| | 1.0 | .963 | .981 | .299 | .150 | .428 | .181 | .036 | .051 |
| Half-Normal | 0 | .052 | .050 | .036 | .014 | .163 | .051 | .001 | .001 |
| | .5 | .528 | .590 | .116 | .041 | .271 | .088 | .007 | .007 |
| | 1.0 | .961 | .970 | .282 | .106 | .425 | .149 | .035 | .028 |
| Logistic | 0 | .047 | .043 | .052 | .066 | .191 | .118 | .002 | .014 |
| | .5 | .534 | .554 | .147 | .171 | .316 | .209 | .013 | .064 |
| | 1.0 | .961 | .970 | .332 | .366 | .463 | .331 | .063 | .198 |
| Cauchy | 0 | .030 | .045 | .029 | .056 | .164 | .103 | .016 | .020 |
| | .5 | .060 | .232 | .031 | .108 | .177 | .139 | .019 | .046 |
| | 1.0 | .125 | .560 | .060 | .195 | .202 | .204 | .021 | .097 |
| Mixed-Normal | 0 | .050 | .049 | .475 | .237 | .559 | .247 | .113 | .102 |
| | .5 | .997 | .994 | .949 | .836 | .897 | .650 | .758 | .738 |
| | 1.0 | 1.000 | 1.000 | 1.000 | .997 | .993 | .942 | .996 | .996 |

| r = .20 (correlated observations) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Laplace | 0 | .247 | .369 | .231 | .271 | .307 | .258 | .138 | .232 |
| | .5 | .546 | .590 | .338 | .389 | .401 | .357 | .206 | .335 |
| | 1.0 | .826 | .851 | .451 | .518 | .506 | .462 | .290 | .456 |
| Lognormal | 0 | .195 | .288 | .154 | .122 | .198 | .120 | .103 | .114 |
| | .5 | .618 | .806 | .217 | .168 | .288 | .167 | .137 | .153 |
| | 1.0 | .876 | .943 | .321 | .237 | .410 | .243 | .193 | .214 |
| Uniform | 0 | .280 | .270 | .299 | .243 | .337 | .230 | .192 | .179 |
| | .5 | .543 | .527 | .380 | .318 | .415 | .293 | .270 | .252 |
| | 1.0 | .790 | .776 | .464 | .400 | .497 | .363 | .352 | .335 |
| Extreme value | 0 | .252 | .267 | .236 | .189 | .282 | .181 | .150 | .152 |
| | .5 | .548 | .572 | .313 | .255 | .363 | .242 | .200 | .205 |
| | 1.0 | .816 | .832 | .399 | .325 | .459 | .317 | .277 | .285 |
| Half-Normal | 0 | .260 | .272 | .243 | .176 | .292 | .173 | .157 | .141 |
| | .5 | .549 | .569 | .315 | .233 | .366 | .229 | .214 | .190 |
| | 1.0 | .810 | .825 | .403 | .305 | .449 | .297 | .279 | .251 |
| Logistic | 0 | .256 | .258 | .252 | .262 | .305 | .246 | .149 | .211 |
| | .5 | .548 | .557 | .352 | .365 | .406 | .331 | .227 | .308 |
| | 1.0 | .817 | .824 | 454 | .467 | .497 | .413 | .307 | .401 |
| Cauchy | 0 | .123 | .276 | .127 | .279 | .227 | .272 | .078 | .247 |
| | .5 | .164 | .401 | .142 | .336 | .237 | .301 | .078 | .278 |
| | 1.0 | .215 | .549 | .156 | .380 | .263 | .341 | .089 | .325 |
| Mixed-Normal | 0 | .200 | .262 | .209 | .266 | .289 | .248 | .107 | .214 |
| | .5 | .609 | .761 | .343 | .430 | .422 | .389 | .180 | .362 |
| | 1.0 | .886 | .955 | .496 | .619 | .573 | .555 | .313 | .557 |




**Robust Statistical Tests, Fig. 7** The probability of rejecting $H_0$ by the *t* test and the Wilcoxon-Mann–Whitney test, as a function of the correlation between successive observations, for normal and exponential distributions ($n_1 = n_2 = 25$ and $m_1 - m_2 = 0$)

non-normal distributions that have various degrees of skewness and kurtosis. The table contains data for a more abbreviated range of values of $m_1 - m_2$: of 0,.5, and 1. It is possible to divide the distributions into two classes. The results for the uniform and logistic distributions were quite similar to those for the normal distribution shown in Figs. 1 and 3. On the other hand, the results for the Laplace, lognormal, extreme-value, half-normal, Cauchy, and mixed-normal distributions were similar to those for the exponential distribution in Figs. 2 and 4.

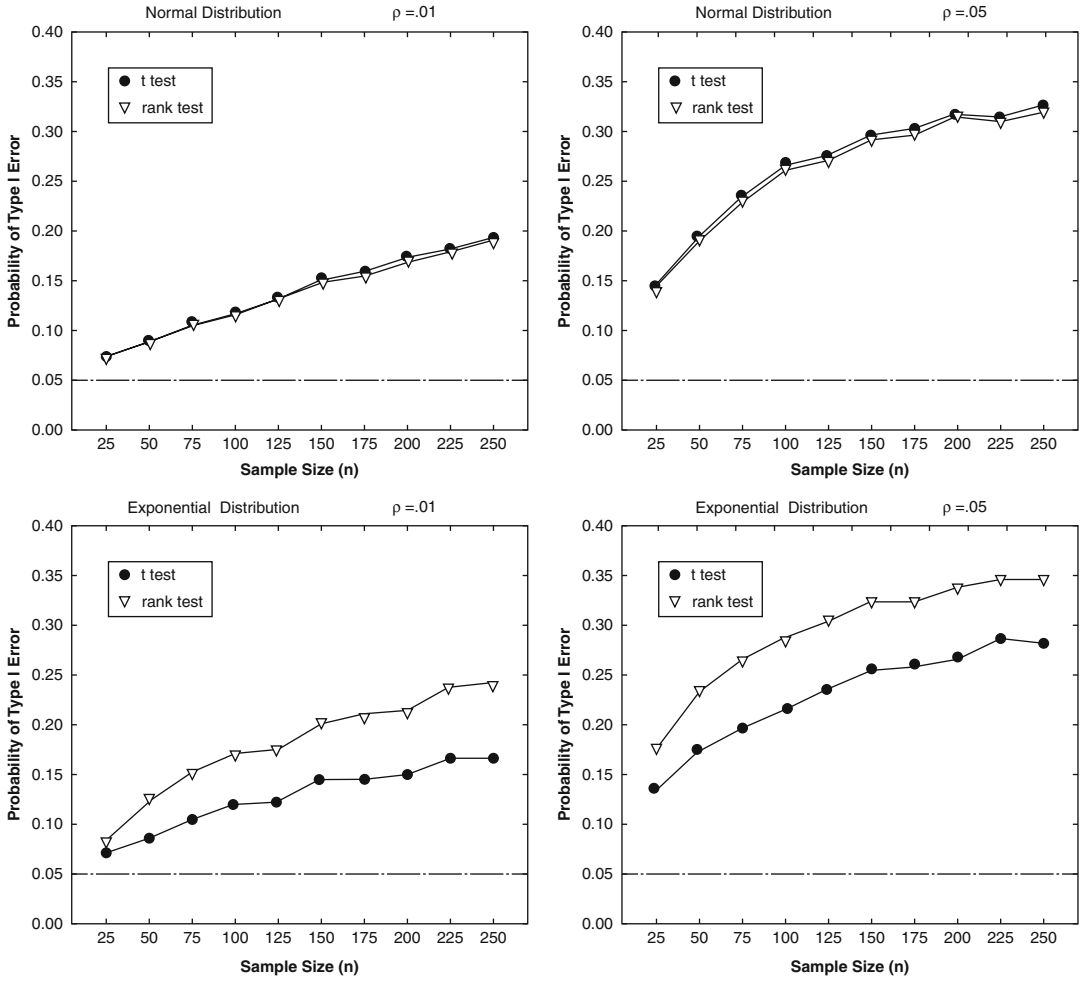
For the cases where sample sizes are equal, standard deviations are equal, and $\rho = 0$, the type I error and power values are similar to those in many studies of the relative power of the *t* and WMW tests based on various non-normal distributions. On the other hand, heterogeneity of variance had pronounced effects for all distributions, even when sample sizes were equal, and the overall results are quite similar to the power functions for the normal and exponential distributions.

When standard deviations were unequal and $\rho = .20$, the results also closely matched the results in the figures for the normal and exponential distributions. In all cases, under heterogeneity of variance, the power of both tests was greatly reduced, and with correlated sample observations, the values were spuriously elevated, so that the curves remained relatively flat.

All these results emphasize that a finding of robustness of the *t* test for one non-normal distribution does not imply robustness for other non-normal distributions. For example, the *t* test is robust for the uniform and logistic distributions, but not for the exponential and lognormal distributions. On the other hand, the WMW test is robust for almost all non-normal distributions, provided that variances are equal and observations are independent.

## Correlated Observations and Sample Size

Violation of independence in sampling has not been studied as much as violations of normality and homogeneity of variance (see, however, Kenny & Judd, 1986; Kruskal, 1988; Stevens, 1996; Zimmerman et al., 1993; Zumbo & Zimmerman, 1991). It turns out, however, that this assumption is crucial in protection of both type I and type II errors, as shown in Figs. 2 and 4. In that case the mutual correlation between sample observations was.30. Figure 7 shows the probability of a type I error for normal and exponential distributions as a function of the correlation. It is evident that the type I error probabilities are severely inflated, even for relatively small correlations.




**Robust Statistical Tests, Fig. 8** The probability of rejecting $H_0$ by the *t* test and the Wilcoxon-Mann–Whitney test, as a function of sample size, when the correlation between sample observations is.01 and.05, for normal and exponential distributions ($m_1 - m_2 = 0$)

R

Figure 8 shows the type I error probability as a function of sample size for the normal and exponential distributions with relatively small degrees of correlation among sample observations of.01 and.05. It is obvious that as sample size increases, the extent of the inflation of the type I error probability systematically increases. It is well known that increasing sample size sometimes leads to a finding of statistical significance for small differences that are not practically significant. The results in this figure indicate that large sample sizes can also magnify the influence of slight degrees of correlation among sample observations that might otherwise be considered inconsequential. The changes are almost the same in the case of the normal distribution, but again the WMW test dominates in the case of the exponential distribution. Clearly, an increase in sample size does not eliminate the effect of correlated observations.

## Summary and Conclusions

It is clear from the above figures and tables that both type I errors and type II errors are relevant to the robustness of a statistical test. Satisfactory control with respect to one does not imply satisfactory control with respect to the other. Not only that, but it is possible for type I error rates to be too high, or

inflated, while at the same time power is spuriously deflated.

Robustness of a significance test when one assumption is violated does not guarantee robustness when another assumption is violated. Some nonparametric methods based on ranks, such as the WMW and Kruskal-Wallis tests, are quite effective in maintaining both type I error rates and power for many non-normal distribution shapes, but they can break down completely, along with *t* and *F* tests, when variances of treatment groups are unequal.

It is well known that the manner in which unequal variances affect type I error rates depends on the sample sizes associated with the respective variances. For normal distributions, if a smaller sample size is associated with a larger variance, the type I error rates of the *t* test are inflated; if a larger sample size is associated with a larger variance, type I error rates are reduced. The same is true for the WMW test, although to a lesser extent. Under normality, type I error rates of the *t* test are not greatly affected when sample sizes are equal, but the type I error rates of the WMW test are affected to a greater extent.

On the other hand, for many non-normal distributions, power is greatly reduced by variance heterogeneity, even if sample sizes are equal, and the reduction occurs at the same time that type I error rates are inflated. This results in the *crossover* effect of the power curves in Fig. 5. It is interesting to note that effects of violation of normality and homogeneity of variance on the probability of rejecting $H_9$ are not additive. That is, if failure of one assumption results in depression of the probability to a certain extent, it is not necessarily true that failure of a second assumption at the same time will result in still further depression. Sometimes the effect is just the reverse, or the degree of the second effect may depend on the degree of the first. Note also that in many cases the power function is relatively flat when variances are unequal. That is, the probability of rejecting $H_0$ does not increase appreciably as the difference between means increases over a range from 0 to 1 s, a result that quite unlike what is ordinarily expected of a significance test.

The effects of violation of independence of sample observations have not been studied widely using simulation methods, but present evidence suggests that such a violation, even if rather small and inconspicuous, can have an substantial effect on type I error rates and power. The present data suggest that such violations always result in a spurious increase in probability of rejecting $H_0$ and apparently are superimposed additively on effects of the other violations. The degree of inflation increases systematically as the mutual correlation between sample observations (the Pearson correlation, ρ) increases. Furthermore, the effect of a fixed degree of correlation becomes increasingly greater as sample size increases.

Considering all these things together, it is clear that the *robustness* of a significance test is a rather complicated attribute that cannot easily be summarized or described briefly. It is difficult to say that a given test is or is not robust under violation of assumptions without considering other parameters. Present knowledge is far more detailed than it was early in the last century, when many investigators often routinely assumed that the *t* and *F* tests are robust to violation of assumptions. It is now evident that the widely used tests based on normal theory cannot be considered robust under *all* possible violations. Furthermore, none of the nonparametric methods based on ranks can be considered effective in counteracting problems that arise in all circumstances. These considerations emphasize the importance of newer procedures, such as bootstrap and resampling methods, that show promise of being effective under a wider range of conditions (see, e.g., Wilcox, 2003, 2005).

## Cross-References

- ▶ Heteroscedasticity
- ▶ *t* Test
- ▶ Type I Error
- ▶ Type II Error

## References

Alexander, R. A., & Govern, D. M. (1994). A new and simpler approximation for ANOVA under variance heterogeneity. *Journal of Educational and Behavioral Statistics, 19*, 91–101.

Blair, R. C., & Higgins, J. J. (1980). A comparison of the power of Wilcoxon's rank-sum statistic to that of Student's *t* statistic under various non-normal distributions. *Journal of Educational Statistics, 5*, 309–335.
Blair, R. C., & Higgins, J. J. (1985). Comparison of the power of the paired samples *t* test to that of Wilcoxon's signed-ranks test under various population shapes. *Psychological Bulletin, 97*, 119–128.
Boneau, C. A. (1960). The effects of violation of assumptions underlying the *t* test. *Psychological Bulletin, 57*, 49–64.
Cressie, N. A. C., & Whitford, H. J. (1986). How to use the two-sample *t*-test. *Biometrical Journal, 28*, 131–148.
Fagerland, M. W., & Sandvik, L. (2009). Performance of five two-sample location tests for skewed distributions with unequal variances. *Contemporary Clinical Trials, 30*, 490–496.
Hodges, J., & Lehmann, E. (1956). The efficiency of some nonparametric competitors of the *t* test. *Annals of Mathematical Statistics, 27*, 324–335.
Hsu, P. L. (1938). Contributions to the theory of Student's *t* test as applied to the problem of two samples. *Statistical Research Memoirs, 2*, 1–24.
Kenny, D. A., & Judd, C. M. (1986). Consequence of violating the independence assumption in analysis of variance. *Psychological Bulletin, 99*, 422–431.
Keselman, H. J., Wilcox, R. R., Taylor, J., & Kowalchuk, R. K. (2000). Tests for mean equality that do not require homogeneity of variances: Do they really work? *Communications in Statistics, Simulation and Computation, 29*, 875–895.
Kruskal, W. (1988). Miracles and statistics: The casual assumption of independence. *Journal of the American Statistical Association, 83*, 929–937.
Randles, R. H., & Wolfe, D. A. (1979). *Introduction to the theory of nonparametric statistics*. New York: Wiley.
Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin, 2*, 110–114.
Scheffé, H. (1970). Practical solutions of the Behrens-Fisher problem. *Journal of the American Statistical Association, 65*, 1501–1508.
Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
Stonehouse, J. M., & Forrester, G. J. (1998). Robustness of the *t* and *U* tests under combined assumption violations. *Journal of Applied Statistics, 25*, 63–74.
Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika, 29*, 350–362.
Welch, B. L. (1947). The generalization of Student's problem when several different population variances are involved. *Biometrika, 34*, 29–35.
Wilcox, R. R. (2003). *Applying contemporary statistical techniques*. New York: Academic Press.
Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing*. New York: Academic Press.
Wolfram, S. (1999). *The mathematica book* (4th ed.). Champaign, IL/New York: Wolfram Media/Cambridge University Press.
Zimmerman, D. W., Williams, R. H., & Zumbo, B. D. (1993). Effect of nonindependence of sample observations on some parametric and nonparametric statistical tests. *Communications in Statistics: Simulation and Computation, 22*(3), 779–789.
Zumbo, B. D., & Zimmerman, D. W. (1991). Further evidence for Coren & Hakstian's "Methodological implications of interaural correlation: Count heads not ears" and an alternative correction formula. *Perception and Psychophysics, 50*, 297–301.

## Robustness

▶ Residuals, Analysis of

## Roland-Morris Disability Questionnaire

Megan Davidson
School of Physiotherapy, La Trobe University, Bundoora, VIC, Australia

### Definition

The Roland-Morris Disability Questionnaire (RMDQ) is a 24-item, low-back-specific functional outcome measure.

### Description

The Roland-Morris Disability Questionnaire (RMDQ) (Roland & Morris, 1983) is a list of 24 statements relating to activities (e.g., walking, dressing, getting up from sitting) and the impairments of pain, appetite, mood, and sleep. Respondents select the statements that apply to them "today." The selected items are summed to a total score ranging from 0 to 24, with a higher score indicating more severe disability. The majority of RMDQ items were derived from the

► Sickness Impact Profile (SIP) and made back-specific by the addition of the words "because of my back." Despite the brevity of the original scale, a number of shorter versions have been published as well as sciatica-specific modified versions.

### Modified Versions of the RMDQ

- A 23-item sciatica version (Patrick et al., 1995) with 19 of the original items. Eleven (Cook et al., 2008) and 12-item short forms (Atlas et al., 2003) of this version
- Sciatica-specific version (Kim, Guilfoyle, Seeley & Laing, 2010)
- 18-item version (Stratford & Binkley, 1997)
- 18-item version (Williams & Myers, 2001)
- 11-item version (Stroud, McKnight & Jensen, 2004)
- Computer-adaptive test version (Cook et al., 2008)

The RMDQ is available in 47 language versions at http://www.rmdq.org/

A yes/no response option is now typically used rather than the original check-box as this allows for differentiation of a "no" response from nonresponse. Online and paper administration has been shown to be equivalent (Bishop et al., 2010).

### Reliability and Measurement Error

Test-retest reliability coefficients for short retest periods (up to 2 weeks) range from $ICC_{2,1} = 0.79$ (Stratford & Binkley, 2000) to ICC = 0.96 (Underwood, Barnett & Vickers, 1999). Longer retest periods have used subgroups who rated their condition as unchanged, with coefficients ranging from $ICC_{2,1} = 0.42$ (Davidson & Keating, 2002) to ICC = 0.93 (Jacob, Baras, Zeev & Epstein, 2001). One study has reported no difference in reliability for the original RMDQ, the 23-item sciatica version, and an 18-item version of the RMDQ (Ostelo, de Vet, Knol & van den Brandt, 2004).

Reports of minimum detectable change (MDC) (90 % confidence) ranges from 4 points (Stratford et al. 1996) to 9 points (Davidson & Keating, 2002). Demoulin, Ostelo, Knottnerus, and Smeets (2010) reported MDC (95 % confidence) ranged from 3.7 points at 1–2 weeks retest to 6.9 points at 12 or more weeks for participants rated as unchanged. The MDC is the magnitude of change score for a stated level of confidence that the change is greater than measurement error.

### Validity and Minimum Clinically Important Difference

Evidence for the construct validity of the Roland-Morris Disability Questionnaire is found in numerous studies correlating scores with measures of performance (e.g., Lee, Simmonds, Novey & Jones, 2001), other disability questionnaires (e.g., Stratford & Binkley, 1997), and known-groups validity (e.g., Leclaire, Blier, Fortin & Proulx, 1997). As would be expected, the Roland-Morris scores have a weaker relationship with the Psychosocial Dimension of its parent questionnaire, the ► SIP, than with the Physical Dimension from which the majority of items are drawn (r = .59 and .89, respectively) (Deyo, 1986).

Rasch analysis shows that RMDQ items cluster in the center of a difficulty continuum, with relatively few items representing easy and difficult activities (Davidson, 2009; Garratt et al. 2003). Comparison the of the 24-item RMDQ, two 18-item, and the 11-item version using Rasch analysis (Davidson, 2009) showed that this problem was more pronounced in the shorter versions where items representing easier and harder activities had been deleted. Two studies (Macedo et al. 2011; Pengel, Refshauge & Maher, 2004) have reported that the 24-item and two 18-item versions have equivalent responsiveness and one (Macedo et al. 2011) that these three versions are more responsive than the 11-item version.

Estimates of the smallest change in scores deemed to be clinically important range from 3.5 points (Ostelo & de Vet, 2005) to 8 points in patients with a high initial score (Stratford, Binkley, Riddle & Guyatt, 1998). Demoulin et al. (2010) reported MCID values increasing with time from 1.5 to 4.5. Jordan, Dunn, Lewis, and Croft (2006) concluded that if RMDQ score improves by 30 % from baseline and the person rates their back pain at least as "better" on

a 6-level global scale, this constitutes clinically important improvement.

### Head-to-Head Comparisons with Other Scales

A Canadian study reported similar reliability and responsiveness of the Roland-Morris, Oswestry, and the Quebec scales, with these scales superior to the SF-36 Physical Functioning Scale (Kopec et al., 1995). An Australian study reported the RMDQ was less reliable than the Oswestry, Quebec, and ▶ SF-36 Physical Functioning scales, but the scales had similar responsiveness (Davidson & Keating, 2002). The RMDQ has been reported as being less responsive than the Patient-Specific Functional Scale, except in samples with severe activity limitation (Hall, Maher, Latimer, Ferreira & Costa, 2011).

One UK study found the RMDQ was overall less responsive than the Oswestry (Frost, Lamb & Stewart-Brown, 2008) while another found it more responsive than the SF-36 or SF-23 in a sample of back-injured workers (Turner, Fulton-Kehoe, Franklin, Wickizer & Wu, 2003). Grotle, Brox, and Vøllestad (2004) reported similar responsiveness of the Norwegian versions of the RMDQ, Oswestry, Disability Rating Index, and SF-36 Physical Functioning scales.

In a post-lumbar surgical population in the Netherlands, Ostelo et al. (2004) reported that three versions of the RDMQ had similar reliability and responsiveness, but these were all superior compared to the SF-36 Physical Functioning and Role Limitations scales.

### Summary

The RMDQ is a brief scale that is simple to complete and score, and its psychometric properties have been extensively researched. There is considerable variability in reports of reliability (range of $MDC_{90}$ from 4 to 9 points) but with 5 points being commonly nominated as representing change beyond measurement error with 90 % confidence. Variability in the MCID is also evident (3.5–8 points). Users should refer to studies with participants similar to their population for the most appropriate values for their purpose.

### Strengths

- Extensive literature on psychometric properties.
- Translations in many languages are available.

### Limitations

- Reference point of "today" may not reflect the sometimes large fluctuations in day-to-day functioning.
- Few items reflect very low- or high-ability activities.
- Lack of content relating to work and social participation.

## References


Atlas, S. J., Deyo, R. A., van den Ancker, M., Singer, D. E., Keller, R. B., & Patrick, D. L. (2003). The Maine-Seattle back questionnaire: A 12-item disability questionnaire for evaluating patients with lumbar sciatica or stenosis: Results of a derivation and validation cohort analysis. *Spine, 28*(16), 1869–1876.

Bishop, F. L., Lewis, G., Harris, S., McKay, N., Prentice, P., Thiel, H., et al. (2010). A within-subjects trial to test the equivalence of online and paper outcome measures: The Roland Morris disability questionnaire. *BMC Musculoskeletal Disorders, 11*, 113.

Cook, K. F., Choi, S. W., Crane, P. K., Deyo, R. A., Johnson, K. L., & Amtmann, D. (2008). Letting the cat out of the bag: Comparing computer adaptive tests and an 11-item short form of the Roland-Morris disability questionnaire. *Spine, 33*, 1378–1383.

Davidson, M. (2009). Rasch analysis of 24, 18 and 11-item versions of the Roland-Morris disability questionnaire. *Quality of Life Research, 18*, 473–481.

Davidson, M., & Keating, J. L. (2002). A comparison of five low back disability questionnaires: Reliability and responsiveness. *Physical Therapy, 82*, 8–24.

Demoulin, C., Ostelo, R., Knottnerus, J. A., & Smeets, R. J. E. M. (2010). What factors influence the measurement properties of the Roland-Morris disability questionnaire? *European Journal of Pain, 14*, 200–206.

Deyo, R. A. (1986). Comparative validity of the sickness impact profile and shorter scales for functional assessment in low-back pain. *Spine, 11*, 951–954.

Frost, H., Lamb, S. E., & Stewart-Brown, S. (2008). Responsiveness of a patient specific outcome measure compared with the Oswestry disability index v2.1 and Roland and Morris disability questionnaire for patients with subacute and chronic low back pain. *Spine, 33*, 2450–2457.

Garratt, A.M., & In collaboration with the United Kingdom Back Pain Exercise and Manipulation Trial. (2003). Rasch analysis of the Roland disability questionnaire. *Spine*, 28, 79–84.

Grotle, M., Brox, J. I., & Vøllestad, N. K. (2004). Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. *Spine, 29*, E492–E501.

Hall, A. M., Maher, C. G., Latimer, J., Ferreira, M. L., & Costa, L. O. P. (2011). The patient-specific functional scale is more responsive than the Roland Morris disability questionnaire when activity limitation is low. *European Spine Journal, 20*, 79–86.

Jacob, T., Baras, M., Zeev, A., & Epstein, L. (2001). Low back pain: Reliability of a set of pain measurement tools. *Archives of Physical Medicine and Rehabilitation, 82*, 735–742.

Jordan, K., Dunn, K. M., Lewis, M., & Croft, P. (2006). A minimal clinically important difference was derived for the Roland-Morris disability questionnaire for low back pain. *Journal of Clinical Epidemiology, 59*, 45–52.

Kim, M., Guilfoyle, M. R., Seeley, H. M., & Laing, R. J. (2010). A modified Roland-Morris disability scale for the assessment of sciatica. *Acta Neurochirurgica, 152*, 1549–1553.

Kopec, J. A., Esdaile, J. M., Abrahamowicz, M., Abenhaim, L., Wood-Dauphinee, S., Lamping, D. L., et al. (1995). The Quebec back pain disability scale. Measurement properties. *Spine, 20*, 341–352.

Leclaire, R., Blier, F., Fortin, L., & Proulx, R. (1997). A cross-sectional study comparing the Oswestry and Roland-Morris functional disability scales in two populations of patients with low back pain of different levels of severity. *Spine, 22*, 68–71.

Lee, C. E., Simmonds, M. J., Novy, D. M., & Jones, S. (2001). Self-reports and clinician-measured physical function among patients with low back pain: A comparison. *Archives of Physical Medicine and Rehabilitation, 82*, 227–231.

Macedo, L. G., Maher, C. G., Latimer, J., Hancock, M. J., Machado, L. A. C., & McAuley, J. H. (2011). Responsiveness of the 24-, 18- and 11-item versions of the Roland Morris disability questionnaire. *European Spine Journal, 20*, 458–463.

Ostelo, R. W. J. G., & de Vet, H. C. W. (2005). Clinically important outcomes in low back pain. *Baillière's Best Practice & Research. Clinical Rheumatology, 19*, 593–607.

Ostelo, R. W. J. G., de Vet, J. C. S., Knol, D. L., & van den Brandt, P. A. (2004). 24-item Roland-Morris disability questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. *Journal of Clinical Epidemiology, 57*, 268–276.

Patrick, D. L., Deyo, R. A., Atlas, S. J., Singer, D. E., Chapin, A., & Keller, R. B. (1995). Assessing health-related quality of life in patients with sciatica. *Spine, 20*, 1899–1908.

Pengel, L. H. M., Refshauge, K. M. & Maher, C. G. (2004). Responsiveness of pain, disability and physical impairment outcomes in patients with low back pain. *Spine*, *29*, 879–883.

Roland, M., & Morris, R. (1983). A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low-back pain. *Spine, 8*, 141–144.

Stratford, P. W., & Binkley, J. M. (1997). Measurement properties of the RM-18. A modified version of the Roland-Morris disability scale. *Spine, 22*, 2416–2421.

Stratford, P. W., & Binkley, J. M. (2000). A comparison study of the back pain functional scale and Roland Morris questionnaire. *Journal of Rheumatology, 27*, 1928–1936.

Stratford, P. W., Binkley, J., Solomon, P., Finch, E., Gill, C., & Moreland, J. (1996). Defining the minimum level of detectable change for the Roland-Morris Questionnaire. *Physical Therapy 76*, 359–365.

Stratford, P. W., Binkley, J. M., Riddle, D. L., & Guyatt, G. H. (1998). Sensitivity to change of the Roland-Morris back pain questionnaire: Part 1. *Physical Therapy, 78*, 1186–1196.

Stroud, M. W., McKnight, P. E., & Jensen, M. P. (2004). Assessment of self-reported physical activity in patients with chronic pain: Development of an abbreviated Roland-Morris disability scale. *The Journal of Pain, 5*, 257–263.

Turner, J. A., Fulton-Kehoe, D., Franklin, G., Wickizer, T. M., & Wu, R. (2003). Comparison of the Roland-Morris disability questionnaire and generic health status measures: A population-based study of workers' compensation back injury claimants. *Spine, 28*, 1061–1067.

Underwood, M. R., Barnett, A. G., & Vickers, M. R. (1999). Evaluation of two time-specific back pain outcome measures. *Spine, 24*, 1104–1112.

Williams, R. M., & Myers, A. M. (2001). Support for a shortened Roland-Morris disability questionnaire for patients with acute low back pain. *Physiotherapy Canada, 53*(1), 60–66.

## Role Balance

▶ Balanced Life

## Role Centrality and Grandparenting

▶ Grandparenting and Psychological Well-Being

## Role Congruity Theory

▶ Job Role Congruence and Gender

## Role Functioning Item Bank

▶ Health and Role Functioning Item Bank Development

## Role Strain

▶ Work-Life Conflict in Europe

## Role Theory

Rosalind C. Barnett
Women's Studies Research Center, Brandeis University, Waltham, MA, USA

### Definition

Role theory conceptualizes everyday activity as the acting out of socially defined categories (e.g., mother, employee, wife). Each social role is a set of rights, duties, expectations, norms, and behaviors that a person has to face and fulfill. This theory is rooted in the observation that people behave in predictable ways and that an individual's behavior is context specific, based on social position and other factors.

### Description

#### Multiple Roles and Quality of Life

At a time when women and men can expect to marry, have children, and be in the paid labor force for most of their adult lives, questions still arise about the impact of multiple roles on ▶ well-being, especially women's. Rarely is the same concern expressed about the deleterious effects of men's multiple roles on men's well-being. Indeed, for men, multiple roles have been seen as beneficial; the better mental health of men compared to women is due in large part to their participation in both family and paid work roles (e.g., Gove & Tudor, 1973). Multiple-role occupancy has been seen as a woman's issue; the role of paid worker is assumed to be added on to the normative condition of being a wife and mother. Role overload, role conflict, guilt, ▶ anxiety, and other hazards are expected to follow, resulting in impaired well-being.

However, merely tallying the number of roles a person occupies fails to acknowledge that particular roles may differ in the extent to which they contribute to well-being. For example, for women, it appears that the worker role is more critical to positive well-being than is the role of parent. For men, adding or participating in the role of father is especially beneficial to well-being. Thus, particular roles may differ in their contribution to well-being, and knowing that a particular role is beneficial to the well-being of one gender does not necessarily mean that it is equally beneficial to the well-being of the other.

#### Dominant Theories

Two major hypotheses have been put forward concerning the relationship between role occupancy and well-being. The "scarcity" hypothesis, put forth by Goode (1960) and others, assumes that the social structure normally creates overly demanding role obligations; these obligations increase as the number of roles one occupies increases.

Because in this view human energy is limited, well-being is impaired by the overload and conflict inherent in numerous, often incompatible roles. In contrast to this view, the "enhancement" hypothesis (Marks, 1977; Sieber, 1974) conceives of human energy as expandable and emphasizes the benefits rather than the costs of multiple-role involvement: status, privileges, increased ▶ self-esteem, and the ability to trade off undesirable components of roles. Like placing one's eggs in many baskets, involvement in several roles is seen as yielding a variety of sources of stimulation, gratification, and social validation.

The weight of empirical research supports the enhancement hypothesis. In general, women and men who engage in multiple roles report lower levels of stress-related mental and physical health

problems and higher levels of subjective well-being than their counterparts who engage in fewer roles (Thoits, 1983, 1986). Although merely occupying a role in no way guarantees that the role will be rewarding, not occupying a role clearly preludes deriving any benefits from it.

### Number of Roles vs. Role Quality

The enhancement and scarcity hypotheses are both limited in that they focus solely on the number of roles occupied; both fail to examine how the nature of a particular role might contribute to or impair well-being. The expansion hypothesis, in contrast, assumes a net gain of benefits over costs regardless of which roles a person occupies. It may be, however, that the particular roles occupied, and the quality of experience in each role, affect the level of well-being more than the mere number of roles. The privileges and obligations, the rewards and concerns, and the cost/benefit balance for a woman who occupies the two roles of wife and mother may differ from those for a woman who occupies the two roles of wife and paid worker. Moreover, two women may occupy similar roles yet experience the quality of each role differently. It is clear that roles are not interchangeable in their impact on well-being: we need to examine not only the quantity but the quality of a women's (and men's) experience in specific roles.

Role quality has at least two components: positive (or rewarding) aspects and the negative (or "of concern") aspects. When an incumbent reports high rewards and low concerns in a role, she/he is said to experience positive role balance. In contrast, if she/he reports low rewards and high concerns in a role, then role balance is said to be negative (Barnett & Baruch, 1985; Baruch & Barnett, 1986).

In sum, it is the qualitative rather than the quantitative aspects of a woman's experiences in her social roles that are the best key to understanding her psychological well-being or lack thereof (Barnett & Hyde, 2001). In other words, role quality is a better predictor of well-being and stress outcomes than is the number of roles occupied.

### Multiple Dimensions of Well-Being

It is now recognized that well-being is multidimensional; high life ▶ satisfaction, ▶ self-esteem, ▶ happiness, and ▶ optimism are positive indicators; and high ▶ anxiety, depression, and ▶ stress-related physical symptoms are negative indicators. Occupancy of a particular role may be related to some but not all well-being indicators. Similarly, positive or negative role quality may be related differently to different dimension of well-being. If well-being is conceptualized as having more than one dimension and if the sources of well-being differ for each dimension, involvement in multiple roles may be a prerequisite for well-being. For example, in one study, a satisfying marriage was strongly tied to happiness, satisfaction, and optimism but was not as strongly associated with self-esteem or depression.

Importantly, relations between well-being and involvement in social roles may vary not only for different roles and different dimensions of well-being but for different groups (e.g., low-income women). More research is needed on these relationships.

### Role Conflict and Role Enhancement

Whenever an incumbent occupies multiple roles, there is the possibility of role conflict or role enhancement. Role conflict arises when the demands from two or more roles are such that adequate performance of one role jeopardizes adequate performance of the others. Two types of role conflict are studied: ▶ family-to-work conflict (fwc) and ▶ work-to-family conflict (wfc).

Role enhancement occurs when the experiences in one role improve the linkage between another role and well-being. So, for example, the well-being of a mother who is experiencing problems in her maternal role will be low unless she is also an employee whose job-role quality is positive.

Finally, just as multiple roles provide opportunities for success, they also offer opportunities for failure or frustration, especially in the context of low-wage work, workplace discrimination, and sexual harassment.

## Cross-References

- ▶ Anxiety
- ▶ Balanced Life
- ▶ Family-to-Work Conflict
- ▶ Gender Role Attitudes
- ▶ Happiness
- ▶ Life Satisfaction, Concept of
- ▶ Optimism
- ▶ Physical Well-being
- ▶ Self-Esteem
- ▶ Stress
- ▶ Well-being and Progress Measurement
- ▶ Women's Well-Being
- ▶ Work-to-Family Conflict

## References

Barnett, R. C., & Baruch, G. K. (1985). Women's involvement in multiple roles and psychological distress. *Journal of Personality and Social Psychology, 49*(1), 135–145.

Barnett, R. C., & Hyde, J. S. (2001). Women, men, work, and family: An expansionist theory. *American Psychologist, 56*(10), 781–796.

Baruch, G. K., & Barnett, R. C. (1986). Role quality, multiple role involvement, and psychological well-being in midlife women. *Journal of Personality and Social Psychology, 51*(3), 578–585.

Goode, W. J. (1960). A theory of role strain. *American Sociological Review, 25*, 483–496.

Gove, W. R., & Tudor, J. F. (1973). Adult sex roles and mental illness. *The American Journal of Sociology, 78*(4), 812–835.

Marks, S. R. (1977). Multiple roles and role strain: Some notes on human energy, time and commitment. *American Sociological Review, 42*, 921–936.

Sieber, S. D. (1974). Toward a theory of role accumulation. *American Sociological Review, 39*, 567–578.

Thoits, P. A. (1983). Multiple identities and psychological well-being: A reformulation and test of the social isolation hypothesis. *American Sociological Review, 48*, 174–187.

Thoits, P. A. (1986). Multiple identities: Examining gender and marital status differences in distress. *American Sociological Review, 51*, 259–272.

# Role-, Relationship-, or Group-Oriented Selves

- ▶ Independent/Interdependent Self

# Romania, Quality of Life

Catalin Zamfir
Romanian Academy, Institute for Quality of Life Research, Bucharest, Romania

## Description

The launch of the ▶ *quality of life* (*QOL*) theme in Western countries did not enjoy a special interest in the socialist countries. Romania was the exception, due to a political-ideological and scientific complex. Here, the of a political and ideological tolerance shown at a special moment in the Romanian communism history. At the same time, the QOL has been assimilated into the public and political language.

### Assimilation of QOL Subject: 1973–1989

QOL subject has been adopted in Romania from the moment it was introduced in the Western countries, by two distinct intellectual communities:

(a) Rapidly, as essays/papers in journals and collective books, by the community of philosophers, specialists in political sciences and economics

(b) By the sociological community, as theoretical papers and as empirical research with a much higher complexity

#### QOL Publications

In this period of time, there were quite a lot of publications, focused explicitly on the QOL.

The first paper which had the QOL in its title was published in 1973, and it was followed by 1–2 studies with such profile each year.

By 1989, *52* such *articles* were published in academic journals.

*Seven books* have been published during this period, with the QOL in their title.

*Five of the* books were collecting distinct studies focusing on QOL subjects:

- A volume with *10* studies by Romanian authors, all focusing to associate the different economic and social-political topics to QOL subject (Institutul de economia comerţului interior şi a turismului 1977)

- A collection of translations from the Western literature and from the socialist countries on QOL (*16* studies) and lifestyle subjects (C. Zamfir & Lotreanu, 1980)
- A volume with several relevant papers (Rebedeu & Zamfir, 1982)
- A *Treaty of contemporary economy* (Postolache, 1987) with *five* chapters dedicated to QOL topics
- A collection of studies on lifestyles (C. Zamfir & Rebedeu, 1989)

*Two of the books* were based on large *empirical researches*:

- *Human development of organizations* (C. Zamfir, 1980a)
- *Indicators and sources of variation of the quality of life* (C. Zamfir et al., 1984)

This last book had an important impact on the academic community, as shown not just by the fact that it was a sample of living, responsible sociology within the bleak ideological context of the time but also by the theoretical and methodological developments and conclusions derived from empirical analysis it contained.

After 1974 a group of sociologist realized *three large empirical researches* on the QOL and quality of work and employment:

- *Human development of organizations* (1976–1977), large samples from six important enterprises (C. Zamfir, 1980)
- *Quality of working life* (*1978–1979*) on a sample of 906 subjects from 16 enterprises (C. Zamfir, Popescu, & Stefanescu, 1980)
- *Diagnosis of quality of life* (1779–1780), on a sample of 1,804 subjects, published in a book (C. Zamfir et al., 1984)

## Political-Ideological and Scientific Context

Why in Romania, unlike other socialist countries, the QOL research and policy subject enjoyed a broad interest among the intellectual community and, furthermore, a tacit political support?

After 1964, Romania wanted to assert its independence from the Soviet Union. The most spectacular gesture was the Romanian decision not to take part to the Soviet intervention, backed by some other socialist states, in Czechoslovakia in 1968. Complementary to the political independence, some kind of ideological independence from the Marxist-Leninist paradigm was asserted. As indirect effect, there was a loosening of the ideological pressure. The ideological control, often formal, was restrained to the mere acceptance of the party policy.

The political and scientific discourse started to be more and more decoupled from the Marxist paradigm. The ideological purity (see ▶ Ideology), the conformity with the classical texts of Marx, Engels, and, particularly, with Lenin, was loosened. In the political and ideological discourse, "Leninism" was cut out from the *Marxism-Leninism* formula, remaining only *Marxism*. The Marxist theory, uncoupled from the Soviet interpretation, became a theoretical corpus with a low degree of coherence. It was increasingly allowed to question the basics of the Marxist paradigm. The ideological confusion was also characterized by a higher tolerance towards new subjects. The initial formulation of *ideological criticism* of the Western intellectual production was replaced by the *critical valorization*. Within this new context, in 1960–1980 there was a large import of Western research and policy issues. In sociology the assimilation of the Western sociological issues and methodology was widely liberalized.

The introduction of QOL issue in Romania occurred thus within the context of chronic crisis of the communist program and of the political closing, but in Romania an increasing ideological confusion.

The signal for the political change was given by the visits of President Ceausescu to North Korea and China in 1973. Repressive actions marked the political change. The Institute of Psychology of the Romanian Academy was closed due to political misunderstanding. Seemingly, it all started from the wrong perception that the psychologists were doing research on the people's evaluation of the political leadership. In fact they confused them with the sociologists. Sociology was marginalized too, the university programs in this specialization being discontinued in 1978. Eventually, sociology was maintained, strongly politicized, and moved to the party university that was training political activists, not sociologists. No empirical sociological research was supported, but some were formally

allowed. Mainly as an effect of discontinuing of the sociology programs at the universities, the empirical programs conducted by the university staff and students had ceased. The empirical researches were performed not within institutional frameworks, but upon the private initiative of some groups of sociologists which took advantage of some political tolerance and of the wide public support. This was the case of the empirical QOL researches by the end of 1970s: social development of organizations, quality of working life, and, mainly, quality of life. In the 1980s, practically no empirical sociological researches were conducted, at least of significant extent.

### QOL as an Alternative Program of Reform

The progressive crisis of the Romanian communist system caused a brutal political closing. Any attempt for structural reform of the system was repressed. There was an unavoidable political deterrence of the sociological researches that had boomed in the 1960s. In this context, contesting the communist model did not seem to have a chance in the predictable future. Without any hope that structural reforms could be possible to promote any more, it was considered, however, that punctual initiatives could lead to or at least prepare global future changes. Thus, on the context of a relative ideological tolerance, some reform programs coming from intellectual environments were developed.

In the second part of 1970s, a group of sociologists launched a program which could be called *the humanization of the social relations and systems* or *improving the quality of life in all social relations and social systems*. The program had two main projects: *quality of working life/human development of organization* and *quality of life*.

The QOL subject in Romania was not just a mere direction of sociological research but mainly the substance of a *sociological program of social reform*. The QOL subject represented a strategy of *meliorism* that might lead, cumulated to other punctual changes, to the change of the whole system or at least to the preparation of such change. It expressed a hope and, in the same time, a pressure for changes of the existing system in some areas where positive effects could eventually be achieved.

The quality of life project had a larger impact keeping alive the pressure for human standards, even in the context of the crisis of the communist model.

### Quality of Life Program

Besides a reform program, the QOL theme had been, in the same time, an answer to the crisis of the Romanian sociology. As of the 1970s, it was increasingly clear that the process of modernization by industrialization and urbanization can no longer be uncoupled from the party policy. The failure of the communist modernization had become more and more apparent. Many subjects of sociological research, dominant in the 1960s, began to lose relevance under the conditions of the generalized crisis of the communist program. Topics such as urbanization, industrialization, and migration from village to towns, in the center of investigation for the sociological research in the previous decade, lost gradually their interest. A void of subjects appeared in sociology. New research topics had to be proposed, which are sufficiently interesting and politically tolerated.

In this context, launching a new program having at its core the concept of the QOL got an attractive opportunity.

The Romanian sociological program of the QOL started from the research of a group from the ▶ Institute for Social Research Michigan (ISR). During the early 1970s, the specialists from that ISR had developed a sociological paradigm of the quality of life and the quality of working life. Complimentarily, they introduced a complex instrument measuring the subjective quality of life, consisting mainly of two groups of indicators: *indicators of satisfaction with life* and *indicators of perceived quality of life*.

**Human Development of the Organizations** The project promoted a new philosophy and methodology in support of the organizations, developed by ISR, based on Likert's theory. The project had two underlying principles. Work satisfaction is an important indicator of organizational performance and an objective by itself of the quality of life. The second principle referred to the democracy at the work place. The intention

was to develop an instrument for diagnosing the social system of the organizations and for changing it (Zamfir & Lotreanu, 1980). A *popularized* version of the project was published as instrument to be used by the programs for human development of the organizations: *A sociologist about work and satisfaction* (C. Zamfir, 1980). The book was also translated in the Soviet Union in an amazing amount of 200,000 copies. Some information suggests that the Russian translation of the book, during a period when sociology was marginalized in the Soviet Union too, enjoyed an unexpected audience.

In the second half of the 1970s, the sociologists still had the illusion that they might support the organizations interested in a process of human development. The contact with the complicated reality of the enterprises turned hopeless such an approach. It was clear from the early 1980s that the only thing to do was to spread, within the university environment, new social patterns of work organization. The Polytechnic Institute of Bucharest set up a department of industrial sociology, which introduced lectures on the *human relations* and *new social patterns* of organizations. Hundreds of students attended such lectures in the 1980s.

In the meantime, the programs of human development of the organizations had been dropped in the Western countries too, under the impact of the neoliberal strategy, and were partially included in the new managerial vision.

In Romania, after the 1989 Revolution, in the process of the chaotic privatization, specific to many post-communist countries, the dream of promoting the quality of working life and democracy at work lost any meaning.

**Quality of Working Life** As a part of the QOL project, an empirical research on the *quality of working life* has been also performed on a sample of 804 subjects, being inspired mainly by the work of Stanley Seashore. The results have been published as an article in a journal with large circulation (C. Zamfir et al., 1980).

**Quality of Life** Several theoretical and methodological studies have been published during the 1970s and 1980s, which promoted the subject and philosophy of the QOL. But the most important impact of the program is due to the empirical research on the QOL realized in 1979–1980. All research was done with private resources but with the important support of the people from the public administration. The initial intention was to make a diagnosis of the quality of life on a representative national sample, but the size of the sample was gradually narrowed to an urban sample, rather large, which could be used for various types of analyses: 1,804 subjects (C. Zamfir et al., 1984).

A large array of perceived quality of life and life satisfaction indicators was incorporated in the survey: global life satisfaction, perceived quality of health, family, job, education, house, environment, incomes, potable water, local community, police activity, mass media, transportation, and leisure opportunities. Moreover, optimism, happiness, and trust in people are included.

Two influential books were published in line with the American humanistic psychology, following the same direction of human reform of the social relations and of promoting the culture of the QOL (E. Zamfir, 1982, 1989).

By that time, the sociological program of quality of life was very popular, being an example of what sociology could and should do.

Within the deep crisis of the Romanian society in the 1980s, the QOL subject remained a mere important scientific, political, and moral gesture. Under the conditions of the social and political crisis of that time, there was a decrease of the publications on the QOL topic. What had to be said was said, but it remained in the collective memory; it seemed useless to continue the effort. However, due rather to inertia, several books and studies prepared previously were published.

## The 1989 Revolution and the Fresh Start of QOL Research

The Revolution of December 1989, as any enthusiastic social change, needed new topics in its program. The QOL was a very attractive subject by that time, being already quite notorious before 1989. The return of the QOL to

the public attention was due to Professor Tudorel Postolache, reputed economist, who played a key role in the development of the reform program of the new government in 1990.

One of the first institutions established after the Revolution was the *Research Institute for Quality of Life*, on January 1, 1990, with 86 employees since the very beginning. As the initiator of the program of quality of life developed in 1970s, the author was appointed director of the new research institute.

Quality of life language became frequently used in the public discourse and even in the institutional construction. After the elections of June 1990, a special position of vice prime minister in charge with the quality of life was introduced.

While the quality of life was the objective of a sociological program of reform during the communist period, after the revolution it was proposed as one of the goals of social reconstruction.

The new *Research Institute for Quality of Life* (RIQOL) was confronted since the beginning with public expectations that exceeded the narrow "classical" sociological QOL paradigm. The Institute tried to integrate the quality of life subject within the complex program of social development. From its establishment, RIQOL aimed not just to make diagnoses of quality of life but also to develop a complex set of research development objectives:

1. The *QOL* in the standard meaning, the set of social indicators, complemented with other socioeconomic indicators: standard of living, incomes, consumption, and lifestyles. The institute decided to make each year the quality of life diagnosis on a representative national sample (see ▶ Quality of Life Diagnosis (QoLD)). Such diagnoses were done annually from 1990 to 1999. Afterwards, the diagnoses were made each 3 years: 2003, 2006, and 2010. After 1989 the Institute produced 13 annual QOL diagnoses. The data for 1990–2010 are included in a consolidated database available to the interested specialists.
2. *Social policies*: what the state does/should do for the society. The expected result was that the quality of life to become a guiding principle of social policies. The QOL topics have been associated to the social policies.
3. *Social problems*: QOL research thematic was diversified with the problems confronting the Romanian society: poverty, social exclusion, poor/abandoned children, and the Roma/Gypsies population.
4. *Programs of social change*: methodology of developing/implementing and evaluating the programs of social change.

After 1989, we can register a spectacular increase in the number of Romanian publications focusing on the QOL, most of them produced by the RIQOL:

- *Papers* with QOL in the title: *135*, with an annual average of 6.4 papers. During the early post-revolution years (1990–1992), there was a burst of publications aiming to draw many specialists towards QOL issues: *39 papers* having QOL mentioned in their title, with an annual average of *13* studies.
- *Books* focusing on QOL topics: *10* (see Institutul de economia comerţului interior şi a turismului 1990; Marginean, 2004; Marginean & Balasa, 2004; Marginean, 2006; Baltatescu, 2009; Marginean & Precupetu, 2010; Mihailescu, 2010; Marginean & Precupetu, 2011).
- *Social reports*. Besides the books and papers intended for a rather small audience, RIQOL specialists also initiated a series of *social reports* which monitor the QOL, analyze critical social problems of the Romanian society, and evaluate the social policies and programs.
- As of 1990, RIQOL publishes a dedicated academic journal: *Quality of life. Journal of social policies*, 4 issues each year. A total of 90 issues have been published in 23 years of its existence.
- Since 2008, RIQOL publishes *Social innovation*, an online open access journal, which enjoys a special audience.

In the context of the Romanian transition, RIQOL, benefiting from its institutional position, has created a substantial enlargement of the quality of life thematic, placing it in a central position in programs of social development.

## Cross-References

- ▶ Communist Regimes, Quality of Life in
- ▶ Ideology
- ▶ Institute for Social Research Michigan
- ▶ Quality of Life
- ▶ Quality of Life Diagnosis (QoLD)

## References


Institutul de economia comerţului interior şi a turismului (1977). *Calitatea vietii* [*Quality of Life*]. *10 papers on quality of life* (*Romanian authors*). Bucuresti: Institutul de economia comertului interior si a turismului.

Institutul de economia comerţului interior şi a turismului (Ed.). (1990). *Calitatea vietii* [*Quality of Life*]. *Proceedings of the Scientific Conference*, *Institute for Quality of Life Research*, *December 13–14* [*Quality of Life*]. Bucuresti: Centrul de informare si documentare economica.

Baltatescu, S. (2009). *Fericirea in contextul social al tranzitiei postcomuniste din Romania [Happiness in the social context of Romanian post-communist transition]*. Oradea: Editura Universitatii din Oradea.

Marginean, I. (2004). *Quality of life in Romania*. Bucuresti: Expert.

Marginean, I. (Ed.). (2006). *First European quality of life survey: Quality of life in Bulgaria and Romania*. Luxembourg: European Foundation for Improving Living and Working Conditions, Office for Official Publications of the European Communities.

Marginean, I., & Balasa, A. (Eds.). (2004). *Calitatea vietii in Romania [Quality of life in Romania]*. Bucuresti: Expert.

Marginean, I., & Precupetu, I. (Eds.). (2010). *Diagnoza anuala a calitatii vietii [Annual Diagnosis of Quality of Life]*. Bucuresti: Expert.

Marginean, I., & Precupetu, I. (Eds.). (2011). *Paradigma calitatii vietii [The paradigm of quality of life]*. Bucuresti: Editura Academiei Romane.

Mihailescu, A. (2010). *Minimul de trai si costurile sociale. Concepte operationale in analiza calitatii vietii [The minimal standard of life and the social costs: Operational concepts in quality of life analysis]*. Iasi: Polirom.

Postolache, T. (Ed.). (1987). *Tratat de economie contemporana [Treaty of contemporary economy]* (Vol. 2). Bucuresti: Editura Politica. *Four chapters dedicated to various aspects of the QOL*

Rebedeu, I., & Zamfir, C. (1982). *Modul de viata si calitatea vietii*. Bucuresti: Editura Politica.

Zamfir, C. (1980a). *Un sociolog despre munca? s?i satisfact?ie*. Bucharest: Editura Politica?.

Zamfir, C. (Ed.). (1980b). *Dezvoltarea umana a intreprinderii*. Bucuresti: Editura Academiei R.S.R.

Zamfir, E. (1982). *Cultura relatiilor interpersonale [Culture of interpersonal relations] (Dezbateri ideologice)*. Bucharest: Editura Politica?.

Zamfir, E. (1989). *Incursiune i?n universul uman: noi ipostaze s?i dimensiuni ale fericirii [Exploring Human Univers. Dimensions of Happiness]*. Bucures?ti: Editura Albatros.

Zamfir, C., & Lotreanu, N. (Eds.). (1980). *Calitatea vietii* [*Quality of Life*]. *10 papers on quality of life* (*Romanian authors*). Bucuresti: 'Stefan Gheorghiu' Academy. *Besides a consistent introductive study, there also were translations of 16 studies from Eastern and Western countries*.

Zamfir, C., Popescu, I.-A., & Stefanescu, S. (1980). Umanizarea muncii – factor al calitatii vietii [Humanizing the work – factor of the quality of life]. *Era socialista, 7*.

Zamfir, C., Popescu, I.-A., Stefanescu, S., Teodorescu, A., Vlasceanu, L., & Zamfir, E. (1984). *Indicatori si surse de variatie a calitatii vietii*. Bucuresti: Ed. Academiei RSR.

Zamfir, C., & Rebedeu, I. (1989). *Stiluri de viata. Dinamica lor in societatea contemporana*. Bucuresti: Editura Academiei.


# Romanian-Hungarian Cross-Border Region, Personal Well-Being Index


Sergiu Baltatescu
Department of Sociology and Social Work and Philosophy, University of Oradea, Oradea, Romania


## Synonyms

Domain satisfaction in Romanian-Hungarian cross-border region; Satisfaction with life as a whole; Well-being, subjective

## Definition

Personal Well-Being Index (PWI) is a scale comprising eight items, namely, satisfaction with standard of living, achievement in life, personal ▶ health, personal relationship, personal safety, community connectedness, future security, and ▶ spirituality. It intends to be a decomposition of ▶ satisfaction with life as a whole. National Well-Being Index (NWI) is a decomposition of "satisfaction with life in the county." The scale

contains six items, i.e., satisfaction with economic situation, state of the environment, social conditions, government, business, and national security. Both scales have been developed by the International Well-Being Group.

## Description

### Overview

In this entry we will report the findings of two surveys administering the PWI and NWI in the North-Western region of Romania (Bihor and Satu Mare counties) and in the neighboring region of Hungary (Hajdú-Bihar and Szabolcs-Szatmár counties). The PWI and NWI showed good psychometric properties (► reliability and validity) consistent with the studies in other countries. Particularities concerning levels in the two countries and relationships with satisfaction with life as a whole are shown and discussed.

### Subjective Well-Being Levels in Romania and Hungary

► Romania and Hungary are among the Central and Eastern European (CEE) countries which have undergone a transition from communist regimes to the capitalist open society. These two countries are among those with low levels of ► quality of life in Europe. This would not be a surprise, given the communist heritage and the economic difficulties which intervened after the regime change. Data from ► World Database of Happiness show that various measures of ► subjective well-being have generally higher values in Hungary than in Romania. For the 4-point life satisfaction scale, used in the Eurobarometer surveys in Romania since 1998 and in Hungary since 2001, the average for Hungary is 50 (transformed on a scale from 0 to 100), while the average for Romania is 44.

### Results of the First PWI and NWI Survey in Romania (2003)

#### Methodology

**Selection of Participants** A representative sample at Bihor county level of around 400 adults was surveyed. The interviews took place at the subjects' homes in November and December 2003. People from 16 localities were interviewed. The selection of households has been realized by a random route. A person aged 18 or more was selected randomly within the chosen household. Non-response rate was around 30 %. About 6 % of the subjects (mainly low-educated rural elders) could not provide reliable answers for well-being items. Two percent of cases were also removed for various reasons, leaving a total number of 368 respondents in the sample. The sample was weighted by sex, age, type of residence, and ethnicity. Finally, Romanians represented 67.8 % of the sample, while Hungarians 30.8 %. Roma, an ethnic group with poorly educated members, which have bad living condition and are generally socially excluded, was underrepresented in the sample. Presumably, the absence of these cases from the sample slightly increased the measured levels on all indicators.

**Characteristics of the Sample** Table 1 displays the characteristics of the samples from 2003 to 2010 (see below).

Along with Personal Well-Being Index (the original seven-item scale version), in the field research which took place in 2003, we also measured the National Well-Being Index. All items have been measured on the recommended 0–10 scale, which, as mentioned before, proved to be difficult to use with the poorly educated rural elders.

#### Personal Well-Being Index Levels

Table 2 summarizes the means, standard deviations, and number of cases for the Personal Well-Being Index using the whole sample.

As we expected, the levels of PWI and life as a whole are substantially lower than those measured in countries such as Australia (see ► Australia, Personal Well-Being Index) but higher than in less economically developed countries like Algeria (see ► Algeria, Personal Well-Being Index).

The highest average rating has the item "personal relationships," which is consistent with results in other countries. Rating of "community connectedness" comes immediately below. PWI

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 1** Demographic characteristics of the sample (difference to 100 % in columns represents the missing values)

| | | | Hungary and Romania (2010) | | |
|---|---|---|---|---|---|
| | | Romania (2003) | Hungarian in Hungary | Hungarian in Romania | Romanian in Romania |
| Gender | Male | 48.2 | 53.2 | 40.1 | 51.5 |
| | Female | 51.8 | 46.8 | 59.9 | 48.5 |
| Age | 18–34 | 36.4 | 19.2 | 23.6 | 27.5 |
| | 35–54 | 34.1 | 50.2 | 32.5 | 49.6 |
| | 55+ | 29.5 | 30.6 | 43.9 | 22.9 |
| Education | Primary or secondary | 32.2 | 24.8 | 27.3 | 24.3 |
| | Vocational | 19.0 | 37.0 | 26.0 | 15.0 |
| | High school | 36.9 | 36.3 | 35.5 | 38.7 |
| | University | 12.0 | 1.9 | 11.2 | 21.9 |
| Marital status | Single | 16.8 | 18.5 | 15.5 | 15.7 |
| | Married | 68.9 | 48.6 | 57.6 | 64.0 |
| | Widowed | 3.3 | 16.7 | 17.1 | 10.0 |
| | Divorced | 9.2 | 11.0 | 4.9 | 6.3 |
| | Lives with partner | 1.0 | 5.2 | 4.9 | 3.9 |
| Residence type | Urban | 51.4 | 63.8 | 67.1 | 59.7 |
| | Rural | 48.6 | 36.2 | 32.9 | 40.3 |

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 2** Means and standard deviations for the PWI items

| Question | Mean | SD |
|---|---|---|
| Personal Well-Being Index | 69.6 | 16.4 |
| Standard of living | 58.7 | 21.0 |
| Health | 66.7 | 26.7 |
| Achievements in life | 64.3 | 23.5 |
| Personal relationships | 80.2 | 18.8 |
| Personal safety | 72.2 | 21.6 |
| Community connectedness | 77.4 | 20.2 |
| Future security | 67.8 | 21.8 |
| Life as a whole | 63.7 | 21.4 |

average level (69.6) is higher than the "life as a whole" average level (63.7). Cummins, Eckersley, Pallant, Van Vugt, and Misajon (2003b, p. 174) found in Australia that, on the contrary, the "PWI" values fall below the level of "life as a whole." They explain this relationship by the fact that the personal index domains are more specific than "life as a whole" and are influenced less by ▶ homeostasis. The Romanian results, similar to those from Algeria (Tiliouine, Cummins, & Davern, 2006), suggest that there is a more complex relationship here. An alternative explanation is based on the fact that PWI is a deconstruction of "life as a whole." As we will show elsewhere in this entry, half of the PWI items have a very low influence on global well-being. This will imply that one or more domain satisfaction, not included in PWI, diminish the global well-being indicator.

### National Well-Being Index Levels

The value for NWI is situated approximately in the middle of the domain values. "National security" has the highest mean (64.4), while "satisfaction with government" the lowest with only 41.8. This last result is not a surprise, as people's dissatisfaction with government is a common characteristic of nations nowadays. "Economic situation" and "social conditions" scores are also very low with around 45 points (Table 3).

NWI items (except "national security") have lower mean scores than each PWI item. This is consistent with other findings about PWI and is explained by Cummins et al. (2003b) by the fact

that domains closer to the respondent are rated higher.

In an international comparison, NWI ratings in Romania situates also, from this point of view, below more economically well-off societies such as Australia and Ireland but again above Algeria (see Fig. 1). The only exception is that of "national security," with ratings above all countries in the set.

Another observation is that, unlike the case of PWI, NWI (50.2) is lower than "life in Romania" (54.3). This is in line with findings from other countries with which we compared ratings in Romania.

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 3** Means, standard deviations, and number of cases (weighted counts). Source: 2003 data, personal calculation

| Question | Mean | SD |
|---|---|---|
| National Well-Being Index | 50.2 | 16.2 |
| Economic situation | 45.3 | 21.2 |
| State of the environment | 52.3 | 22.3 |
| Social conditions | 45.2 | 20.2 |
| Government | 41.8 | 25.5 |
| Business | 50.9 | 20.6 |
| National security | 64.7 | 23.0 |
| Life in Romania | 54.3 | 21.9 |

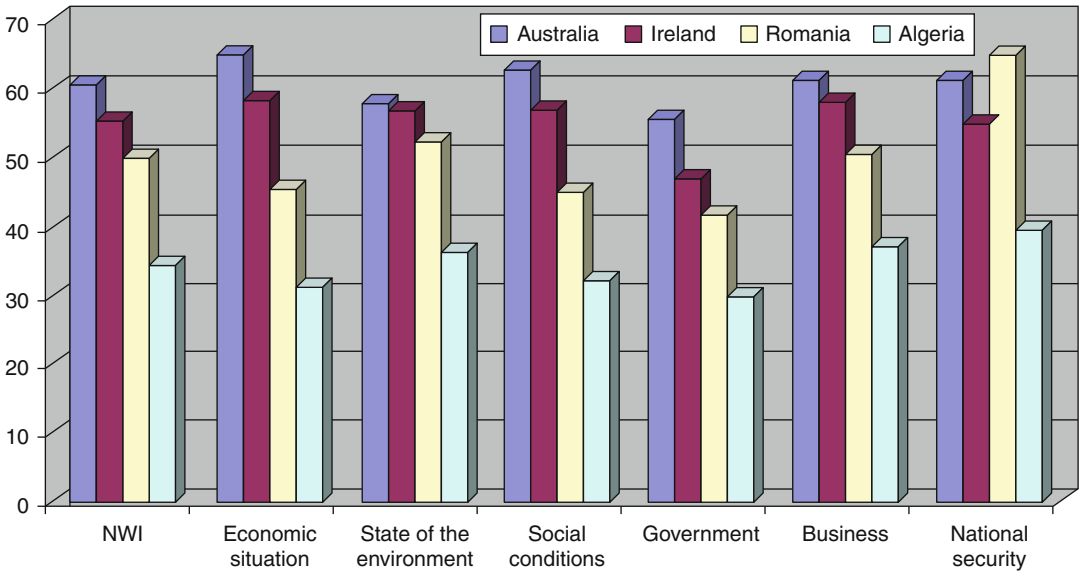
### Psychometric Properties: External Validity and Reliability

Since the PWI is a deconstruction of life as a whole, they must have a high correlation with one another. The same relationship would be expected with NWI and life in Romania. Both items (see Table 4) have higher correlations with similar measures, higher than with each other.

Both indexes show a very good reliability, with Cronbach's alphas over 0.8 (PWI 0.86 and NWI 0.82).

### Factorial structure

All items from both indexes were introduced in a ▶ Principal Component Analysis. The results were rotated using Varimax method with Kaiser Normalization. Three factors emerged, with the first and second capturing most items from PWI and NWI, respectively, and the third explaining some items from both indexes. However, the third factor is weak (with eigenvalue around 1.05).




**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Fig. 1** Comparison between average levels of NWI items for Australia, Ireland, Romania, and Algeria (Sources: 2003 data, personal calculation for Romania, Cummins et al. (2003a) for Australia, and Brownlee and O'Neil (2003) and Tiliouine et al. (2006) for Algeria)

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 4** Bivariate correlations between "life as a whole," PWI, NWI, and "life in Romania." Source: 2003 data, personal calculation. All coefficients are significant at .001 level

| | Personal Well-Being Index | Life as a whole | National Well-Being Index |
|---|---|---|---|
| Life as a whole | .72 | | |
| National Well-Being Index | .41 | .44 | |
| Life in Romania | .47 | .54 | .67 |

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 5** Two-factor solution for Personal and National Well-Being Items. Source: 2003 data, personal calculation

| | Factor 1 Personal WB | Factor 2 National WB |
|---|---|---|
| Life as a whole | .72 | .33 |
| Standard of living | .69 | .34 |
| Health | .62 | .01 |
| Achievements in life | .71 | .24 |
| Personal relationships | .74 | .06 |
| Personal safety | .77 | .20 |
| Community connectedness | .75 | .07 |
| Future security | .74 | .30 |
| Life in Romania | .37 | .71 |
| Economic situation | .19 | .85 |
| State of the environment | .23 | .58 |
| Social conditions | .22 | .81 |
| Government | .00 | .66 |
| Business | .12 | .77 |
| National security | .13 | .61 |
| Variance explained | 41.84 | 14.72 |
| Total variance explained = 56.57 % | | |

When a 2-factor solution was requested, the results clearly dissociate National from Personal items (Table 5).

### Unique Contribution of Domains to Life as a Whole: Multiple Regression

We regressed life as a whole against the seven PWI items in order to compare the unique contribution of the domain satisfaction to the global subjective well-being indicator (see Table 6). The overall fit is very good: variances of PWI items explain 59 % of the variation of the dependent variable. Satisfaction with the standard of living, with beta $\beta = .47$, has the largest contribution to the prediction of life as a whole. Other domains that also predict the independent variable are personal achievements, $\beta = .16$; personal relationships, $\beta = .11$; and future security, $\beta = .11$. Three domains (health, safety, and community connectedness) have nonsignificant independent contributions to life as whole. However, Cummins et al. (2003a) did not favor the idea of excluding these domains from PWI as they predict life as a whole in other countries.

## Results of the Second PWI Survey in Romania and Hungary (2010)

### Methodology

**Selection of Participants** The field research took place within the framework of ENRI (European, National and Regional Identity – Theory and Practice) project, financed by the Hungary-Romania Cross-Border Co-operation Programme 2007–2013. We planned a random stratified sample of around 1,000 subjects in each of the two countries. The data were collected in a multistage approach. Firstly, the localities in which the interviews would take place were selected from four neighboring counties (two in Hungary and two in Romania). The sample of localities was stratified, the settlements being split in three categories: county capitals (large cities), townships in the area, and villages. In the second stage, we selected the subjects. Their number for each settlement was computed based on two criteria: proportionality and a number of at least 30 subjects for each locality. The sample has very good representativeness with the single exception of rural/urban breakdown in Romania where we have an overrepresentation of urban population and with that of education in Hungary, where respondents with university degrees were strongly underrepresented. The final number of interviews was of 1,000 subjects from Hungary and 824 from Romania. More details and a deeper analysis of the results can be found in Baltatescu (2011).

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 6** Bivariate correlations between PWI items and regression coefficients and their significance (dependent variable: "life as a whole"). Source: 2003 data, personal calculation

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | B | b | Sig (t) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Life as a whole | | | | | | | | | | |
| 2. Standard of living | .72 | | | | | | | .48 | .47 | .00 |
| 3. Health | .41 | .37 | | | | | | .04 | .05 | .24 |
| 4. Achievements in life | .58 | .57 | .44 | | | | | .15 | .16 | .00 |
| 5. Personal relationships | .49 | .46 | .30 | .42 | | | | .13 | .11 | .01 |
| 6. How safe you feel | .52 | .52 | .39 | .53 | .50 | | | .01 | .01 | .90 |
| 7. Community connectedness | .48 | .43 | .38 | .46 | .55 | .57 | | .05 | .05 | .31 |
| 8. Future security | .54 | .48 | .41 | .56 | .47 | .71 | .61 | .11 | .11 | .03 |
| Adj $R^2 = .59$ | | | | | | | | | | |

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 7** Ethnic group by country. Source: 2010 data, personal calculation

| Ethnic group | % of the overall sample in the country |
|---|---|
| Romanian in Romania | 70.1 |
| Hungarian in Romania | 29.9 |
| Hungarian in Hungary | 97.9 |

**Characteristics of the Sample** The subjects of Hungarian ethnicity make up 29.9 % of the Romanian sample and 97.9 % of the Hungarian sample (Table 7).

Three main groups will make the comparative focus, based on ethnicity and country: Hungarians in Hungary, Romanians in Romania, and Hungarians in Romania.

While life satisfaction was measured, as recommend by the authors of the Index, by the standard question with 11 points, for the Personal Well-Being Index items, we used a 5-point verbalized scale ranging from "Very Dissatisfied" to "Very Satisfied," considering that the verbalized questions will be more easily understood by the rural population in the areas of the study. The scores obtained on 1–5 scales were linearly translated to the standard 0–10 scale.

### Levels of Personal Well-Being and Life Satisfaction

A comparison of average levels for each item within the three subsamples shows no significant differences for "standard of living" and "health." The ethnic majority in Hungary reports higher satisfaction with "achievements in life" but lower satisfactions with "religion/spirituality" than the ethnic majority in Romania. However, the average levels for these items are not different from those reported by the Hungarians in Romania (Table 8).

For the rest of domains, respondents from Hungary reported higher average levels than those from Romania. Within the Romanian sample, Hungarian minority members report similar levels with Romanian majority, with the exception of community connectedness. Overall, respondents from Hungary report the highest levels of personal well-being, which is perfectly explainable given that, on average, Hungary has better economic and social conditions than Romania, and it also fits the average differences in life satisfaction between the two countries. However, when we measure differences in life satisfaction, we find that on average the respondents from Hungary report the lowest scores.

As a first explanation, we suggest that there may be a national factor that lowers the average answers to the overall life satisfaction in case of Hungarians. This factor, which can be a more depressive mood or a different interpretation of the life satisfaction question, apparently makes the Hungarian ethnics report lower levels of satisfaction with life as a whole, contrary to the highest evaluation of most life domains. The hypothesis of a "national character," although

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 8** Average levels of personal well-being items for the three samples, standard deviations, and sign of significant differences between the means (using Bonferoni test). Source: 2010 data, personal calculation

| | Ethnicity by country | | | | | | Sign of significant differences | | |
|---|---|---|---|---|---|---|---|---|---|
| | Hungarian in Hungary (1) | | Hungarian in Romania (2) | | Romanian in Romania (3) | | | | |
| | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. | 1–2 | 1–3 | 2–3 |
| Standard of living | 31.90 | 9.79 | 31.34 | 9.66 | 32.06 | 9.80 | | | |
| Health | 35.26 | 11.61 | 33.46 | 10.09 | 35.02 | 8.75 | | | |
| Achievements in life | 36.91 | 8.67 | 36.14 | 8.04 | 35.12 | 7.69 | + | | |
| Personal relationships | 40.53 | 8.20 | 38.54 | 6.41 | 37.27 | 6.90 | + | + | |
| Personal safety | 39.61 | 8.58 | 35.57 | 7.69 | 35.42 | 7.78 | + | + | |
| Community connectedness | 38.99 | 9.15 | 37.44 | 6.91 | 35.65 | 7.21 | + | + | – |
| Future security | 34.75 | 10.73 | 31.71 | 8.35 | 31.88 | 9.09 | + | + | |
| Religiousness/spirituality | 37.42 | 9.85 | 38.74 | 7.64 | 38.68 | 6.87 | – | | |
| Personal Well-Being Index | 67.30 | 15.90 | 63.41 | 13.54 | 62.84 | 13.82 | + | + | |
| Life as a whole | 59.35 | 18.90 | 61.95 | 19.88 | 66.21 | 18.47 | | – | – |

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 9** Bivariate correlations between Personal Well-Being Index and life as a whole. The r coefficients are significant at 0.01 level. Source: 2010 data, personal calculation

| | Pearson's R |
|---|---|
| Hungarian in Hungary | .572 |
| Hungarian in Romania | .418 |
| Romanian in Romania | .540 |

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 10** One-factor solution for Personal Well-Being Items. Source: 2010 data, personal calculation

| | Hungarian in Hungary | Hungarian in Romania | Romanian in Romania |
|---|---|---|---|
| Standard of living | .760 | .725 | .753 |
| Health | .724 | .724 | .744 |
| Achievements in life | .723 | .698 | .737 |
| Personal relationships | .710 | .692 | .725 |
| Personal safety | .673 | .646 | .673 |
| Community connectedness | .607 | .637 | .652 |
| Future security | .606 | .636 | .648 |
| Spirituality or religion | .530 | .588 | .572 |
| Total variance explained | 44.98 % | 44.86 % | 44.68 % |

frequent in the literature, is nonetheless highly disputable (Veenhoven, 1984). A second explanation points out that Romanian respondents give higher ratings to life satisfaction scales with 10–11 points. Although average life satisfaction levels are higher in Hungary than in Romania, in the European quality of life survey (EQLS) (2003, 2007), which uses a 0–10 scale, Romanian respondents give consistently higher ratings than Hungarian respondents. This second hypothesis is linked with the "national character" but suggests a methodological bias.

### Psychometric Properties: External Validity and Reliability

As a decomposition of the "life as a whole" variable, PWI should have a relatively high correlation with it. Indeed, the overall bivariate correlation between these variables is 0.50. The Pearson correlation coefficient for the subgroups varies as follows (Table 9):

### Factorial Structure

The eight PWI items have been introduced in a Principal Component Analysis. The results

**Romanian-Hungarian Cross-Border Region, Personal Well-Being Index, Table 11** Regression of personal well-being items against life satisfaction for the three subsamples. Source: 2010 data, personal calculation

| | Hungarian in Hungary | | Hungarian in Romania | | Romanian in Romania | |
|---|---|---|---|---|---|---|
| | β | Sig(t) | β | Sig(t) | β | Sig(t) |
| Standard of living | 0.35 | 0.00 | 0.44 | 0,00 | 0,30 | 0,00 |
| Health | 0.02 | 0.44 | 0.02 | 0,73 | −0,03 | 0,58 |
| Achievements in life | 0.13 | 0.00 | 0.17 | 0,01 | 0,17 | 0,00 |
| Personal relationships | 0.03 | 0.42 | 0.03 | 0,63 | 0,00 | 0,99 |
| Personal safety | 0.16 | 0.00 | −0.09 | 0,18 | 0,11 | 0,02 |
| Community connectedness | 0.04 | 0.20 | −0.04 | 0,60 | 0,08 | 0,07 |
| Future security | 0.08 | 0.01 | 0.13 | 0,06 | 0,02 | 0,71 |
| Spirituality or religion | 0.05 | 0.06 | −0.06 | 0,37 | 0,10 | 0,02 |
| Adj. $R^2$ | .381 | | .321 | | .327 | |

were rotated using Varimax method with Kaiser Normalization. For the subsample of Romanians in Romania, a single factor emerged. For the other two subsamples, a solution with two factors was computed. In the case of the subsample of the Hungarians in Hungary, the second factor explained the following items: "spirituality/religion," "community connectedness," and "personal relationships," while "future security" is loading in both factors almost equally.

For the Hungarian in Romania subsample, the second factor explains the same four items, except that instead of "future security," it explains "achievements in life." However, the second factor is very weak, with eigenvalues very little above 1 (1.018–1.0123). Ordering a one-factor solution, we obtain the results showed in Table 10.

It was also proved that the PWI scale has an excellent reliability for all the subsamples (the Cronbach's alpha varies from 0.815 to 0.836).

### Unique Contribution of Domains to Life as a Whole: Multiple Regressions

To compare the unique contribution of domain satisfaction to the global subjective well-being indicator, we regressed life as a whole against the PWI items. Table 11 shows the beta coefficients for all the subsamples. The predictors explain between 32.1 % and 38.1 % of the variation of the dependent variable.

Among the predictors of life satisfaction, "standard of living" has the largest contribution, with beta coefficients varying from 0.30 to 0.44. The list of predictors include "personal achievements" (β = .13 to .17), "personal safety" (β = −.09 to 0.11), and "future security" (β = .08, .13 and .02 (nonsignificant)). Three domains (health, safety, and community connectedness) have nonsignificant independent contributions to life as whole. Spirituality/religion is a predictor only for the Romanian subsample (β = .10).

## Conclusions

Overall, we can conclude that PWI and NWI have good psychometric properties in the Romanian-Hungarian cross-border region for all the ethnic groups researched. Moreover, membership in these groups influences both the levels and the contribution of specific dimensions to the explanation of "satisfaction with life as a whole." Specifically, for the Romanians, a more religious population, the "religiousness/spirituality" item has a positive contribution to the explanation of life as a whole, which justifies the inclusion of this item in the Index. Respondents with Hungarian ethnicity, on both sides of the border, report being less satisfied with life as a whole, although their average PWI levels are higher than those of Romanians. Given that their level of living is higher, this is a suggestion that Personal Well-Being Index is a better fit to be used as an overall satisfaction indicator than "life as a whole," being less influenced by "national character" or other methodological biases.

## Cross-References

- ▶ Algeria, Personal Well-Being Index
- ▶ Australia, Personal Well-Being Index
- ▶ Communist Regimes, Quality of Life in
- ▶ Domain Satisfaction
- ▶ European Quality of Life Survey (EQLS)
- ▶ Homeostasis
- ▶ Quality of Life
- ▶ Romania, Quality of Life
- ▶ Satisfaction with Life as a Whole
- ▶ Subjective Well-Being
- ▶ Subjective Well-Being Indicators
- ▶ World Database of Happiness

## References

Baltatescu, S. (2011). Subjective well-being in counties at the cross-border between Hungary and Romania. In B. Balogh, K. Bernáth, I. Bujalos, A. Hatos, & I. MurányI (Eds.), *European, national and Regional Identity: Theory and Praxis* (pp. 319–332). Oradea: Partium.

Brownlee, C., & O'Neill, G. (2003). Quality of Life in Ireland. *St. Patrick's Festival Symposium*.

Cummins, R. A., Eckersley, R., Lo, S. K., Okerstrom, E., Hunter, B., & Davern, M. (2003a). *The wellbeing of Australians – The effects of work. Australian Unity Wellbeing Index. Report 7.0*. Melbourne, Victoria: Australian Centre on Quality of Life. Deakin University. http://acqol.deakin.edu.au

Cummins, R. A., Eckersley, R., Pallant, J., Van Vugt, J., & Misajon, R. (2003b). Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. *Social Indicators Research, 64*(2), 159–190.

Tiliouine, H., Cummins, R., & Davern, M. (2006). Measuring wellbeing in developing countries: The case of Algeria. *Social Indicators Research, 75*(1), 1–30. doi:10.1007/s11205-004-2012-2.

Veenhoven, R. (1984). *Conditions of happiness*. Dordrecht, Holland/Boston, Hingham, MA: Kluwer.

## Romantic Desire

- ▶ Love

## Romantic Unions

- ▶ Dating Relationships

## Rooflessness

- ▶ Homelessness

## Rootedness

- ▶ Place-Related Measures

## Roots Music

- ▶ Folk Music

## Routine Activities Theory of Crime and Victimization

- ▶ Victimization and Routine Activities

## Routine Activity Theory

- ▶ Impact of Housing Design on Crime

## RPWB

- ▶ Psychological Well-Being Inventory

## Rubin Causal Model

- ▶ Potential Outcomes

## Rule of Law

- ▶ Trias Politica (Separation of Powers)

## Rural Development Index (RDI)

▶ Rural Development Index Applied to Poland and Slovakia

## Rural Development Index Applied to Poland and Slovakia

Jerzy Michalek[1] and Nana Zarnekow[2]
[1]European Commission, Seville, Spain
[2]Department of Agricultural Economics, University of Kiel, Kiel, Germany

### Synonyms

Rural Development Index (RDI)

### Definition

The rural development index (RDI) is a multidimensional (composite) indicator measuring the overall level of rural ▶ development and the ▶ quality of life (QOL) in individual rural regions. The rural development domains can be represented by numerous partial socioeconomic, environmental, infrastructural, and administrative indicators variables (Michalek & Zarnekow, 2012). The RDI aggregates the above domains into a one-dimensional indicator using *objective* and *statistically verifiable weights*. As a composite indicator (CI), the RDI should fulfil a number of important general conditions (Hagerty et al., 2001; OECD, 2005), e.g., it should be based on a sound theoretical framework; the selection of variables should take into consideration their relevance, accessibility, etc.; construction of the index should follow an exploratory analysis investigating the overall structure of used indicators; and the index should be reported as a single number but could be broken down into components/domains. In order to be relevant for an empirical ▶ policy analysis, a composite QOL/RDI index should also meet a number of general (e.g., efficiency, effectiveness, relevance) and specific (e.g., regionality, rurality, simplicity) policy criteria (Kaufmann, Stagl, Zawalinska, & Michalek, 2007). Given growing demand for composite development indicators in applied policy analysis (e.g., in evaluation of rural development/structural programs), potential gains from having a multidimensional regional/rural development index are straightforward. As a composite indicator, the RDI can be applied to analysis of the main determinants of rural/regional development in individual rural areas (Michalek & Zarnekow, 2012) as well as for the assessment, i.e., the measurement of the impact of rural development and structural or cohesion policy (Michalek, 2012).

### Description

The review of various empirical studies concerned with construction of a composite index to policy analysis shows that its creators have to cope with numerous methodological issues of which the most crucial ones are the following (Kaufmann et al., 2007): (1) selection of appropriate variables/coefficients and balancing between objective vs. subjective indicators, (2) weighting the variables/indicators according to their relative importance, (3) application of unbiased aggregation techniques, and (4) making the index useful for policy purposes (i.e., in ▶ program evaluation). Other major difficulties associated with construction of a relevant composite indicator can be summarized as follows: (1) in a huge majority of relevant studies, the choice/selection of "the most representative" socioeconomic indicators was carried out *arbitrary*, *leaving other available indicators unused or downgraded* as "nonrepresentative"; (2) experts' weights of selected indicators appear often as extremely *subjective* and not directly transferable from one geographic area to another; (3) different normalizations of variables could result in different weights; (4) some weights would become inconsistent when a larger number of indicators/coefficients/variables had been analyzed; (5) weights that were based on a pure statistical analysis of factors (e.g., based on

factor loadings) appeared to miss an appropriate ▶ welfare (social utility) context; and (6) many assigned weights tended to be only region specific, so they are not applicable to other regions even in the same country.

**The rural development index** (**RDI**) developed in Michalek and Zarnekow (2012) directly responds to above challenges and measures multidimensional concepts of rural growth/decline by embracing performance of the most important rural development domains, e.g., economic output (including agriculture, food industry, rural tourism), investment, employment, ▶ poverty, ▶ education, ▶ health, ▶ housing conditions, ▶ crime, environment, urbanization, and land use. The weights of economic, social, and environmental domains entering the RDI index are derived empirically from the econometrically estimated intra- and interregional migration function, which inter alia takes into consideration preferences of both migrants as well as those who stayed, and can be therefore viewed as representative for a whole population in a given time period. The RDI developed in Michalek and Zarnekow (2012) can be applied to analysis of the main determinants of rural/regional development in a given country both at aggregated (NUTS-2) as well as highly disaggregated levels (e.g., NUTS-4); examination of specific economic, social, and environmental factors affecting various aspects of the quality of life and rural development at a regional/local level; measurement of the level in regional disparities in overall regional development (beyond GDP); ranking rural areas with respect to their overall (synthetic) level of development; and a quantitative evaluation (i.e., the measurement of the impact) of specific policies, e.g., rural, structural, and cohesion affecting rural economies.

## Methodology Used for Derivation of the RDI

Assuming equivalence between the level of rural development and rural quality of life, the methodology used in Michalek and Zarnekow (2012) for derivation and construction of a composite RDI draws upon research on the relationship between the quality of life and migration. The original foundation for analyzing the effect of regional performance and migration was provided by Tiebout (1956), who found out that as long as consumers are fully mobile and informed, they convey their preferences through migration or "voting with their feet." A vast sociologic and economic literature shows that people tend to move in order to improve the quality of their lives in a variety of specific respects, and they continue to move until they achieve goals for the majority of those respects (Berger, Blomquist, & Sabirianova, 2003; Douglas & Wall 1993, 2000; Fuguitt, 1985; Michalos, 2003).

In migration studies incorporating characteristics of origin and destination regions, the most frequently reported motives for in-migration flows into destination areas (pull factors) included factors such as higher probability of obtaining employment, better housing, nicer neighborhood, lower pollution, lower crime rates, and better health service. Under factors found to determine out-migration in origin areas (push-factor), the most important were poor location amenities, poor public transportation, lack of good medical facilities, ▶ unemployment, economic and environmental distress, etc. (Greenwood 1997; Michalos, 2003). Furthermore, various migration studies showed empirically that people living in societies that have reached a certain stage of material wealth will also increasingly focus upon immaterial aspects of life, e.g., attractiveness of places that depends upon the needs, demands, and preferences of the individual (Inglehart, 1997; Niedomysl, 2006). The "pull-push" approach assumes that numerous *objective indicators* describing various regions (e.g., unemployment, crime rate, infant mortality, level of prices) can be transformed into a *subjective judgement* of the overall quality of life on which any migration decision is made.

The methodological approach applied in Michalek and Zarnekow (2012) for derivation of *weights* in RDI draws on the supposition that quality of life and migration are closely linked to each other and builds upon Tiebout (1956), Douglas and Wall (1993, 1997), and (2000) who argue that cross-migration rates provide the richest and most reliable source of data on

the relative attractiveness of different locations. Yet, contrary to previous studies, the approach used in this study *neither implies equivalence between quality of life and migration*, nor is the quality of life expressed as a parameter that is *independent of individual characteristics* of a given location. In fact, as we show below, the method proposed in this study allows for the computation of the quality of life/rural development index even in regions exhibiting null in- or out-migration.

From the perspective of this study, particularly interesting versions of migration models are those models which forecast probability of migration by incorporating information on the relative frequency of non-migration (e.g., probit or logit models) thus providing a natural transition from the gravity model to the more behavioral-grounded modified gravity models. The modelling of migration decision depends also on the type of data available. Irrespective of a selected object of such analysis (individual or household) and chosen methodological approach (nonspatial vs. spatial econometrics), major determinants of a migration decision appear those variables describing:

- Differences in factors determining the quality of life in origin and destination regions
- Transaction costs related to such a decision

Formally, the RDI in region *i* can be expressed as a linear function of *i*-region-specific characteristics $\mathbf{Z_{ki}}$ and their weights $\boldsymbol{\beta_k}$:

$$\mathbf{RDI_i} = h\left(\beta_k, Z_k{}^i\right) = \Sigma_k \beta_k * Z_k{}^i \qquad (1)$$

where $\mathbf{RDI_i}$ is the rural development index (an equivalent of the quality of life index) in region i; $Z_k{}^i$ are measurable characteristics k in a region i; and $\boldsymbol{\beta_k}$ are weights for each characteristic $\mathbf{Z_k}$ derived from a given migration model.

In empirical work, due to the multidimensionality of relevant data, a particular importance is to be assigned to:

1. An appropriate selection (or estimation) of $Z_k{}^i$ describing major attributes of the overall development and the quality of life in individual rural areas
2. Appropriate estimation of (social) weights $\boldsymbol{\beta_k}$

The estimation of the RDI in Michalek and Zarnekow (2012) was carried out by taking the following steps:

1. Defining relevant rural development domains to be taken into consideration prior to the assessment of the overall impact of specific RD policies/program.
2. Defining variables describing each rural development domain in all regions **i**.
3. Translating the above variables into meaningful coefficients (e.g., per capita, per $km^2$) in all regions i.
4. Converting those coefficients into region specific factors $\mathbf{f_i}$ (principal component method) in order to reduce the dimension of the analysis (► factor analysis).
5. Deriving *weights* for each individual factor/principal component $\mathbf{f_i}$ (embracing respective variables in each rural development domain) to be applied in the construction of the RDI from econometrically estimated migration function (using a multilevel mixed-effect regression model). It is important to note that introduction of transaction costs into the migration model brings about a formal separation of the RDI (consisting of individual factors and related estimated coefficients) from migration. This is because transaction costs do not enter the index itself but are used to explain a part of the overall variance in a migration model.
6. Computing for each rural region in a synthetic index $\mathbf{RDI_i}$. The latter is defined as a weighted sum of factors (variables, domains) with $\boldsymbol{\beta_k}$ derived from a selected inter- and intra-regional migration function according to Eq. 1 (the optimal number of factors k selected to the construction of an RDI was derived from the maximization of the restricted likelihood function used in the estimation of the intra-regional migration model).

Generally speaking, existing literature does not provide a definite answer to the question: which domains and what relevant variables/proxies should be selected into a synthetic/composite index measuring the overall level of economic and social development/quality of life

(Erikson, 1993; Grasso & Canova 2007; Johansson, 2002; Jones & Riseborough, 2002; Kazana and Kazaklis 2008). In order to meet relevant policy criteria (e.g., objectivity, transparency, and simplicity) and ensure full data comparability across all regions within a given country, an indirect approach was applied in Michalek and Zarnekow (2012). In this approach, a country's available secondary regional statistics (objectively verifiable indicators) representing various aspects of quality of life were used, instead of subjective indicators derived on the base of sporadic interviews with individuals in selected regions (NUTS-4). An important advantage of this approach is an explicit consideration of all aspects of regional/rural development available from secondary statistics at regional basis (i.e., economic, social, environmental, infrastructural, administrative), thus avoiding an arbitrary preselection of "the most important" partial indicators, by using subjective judgments as to their "social relevance" and "representativeness."

Furthermore, the applied method allows for the assessment of "social importance" of all individual partial indicators collected at regional level and their respective domains. Given estimates of $\boldsymbol{\beta}_k$ (<social> weights) for all individual factors $\mathbf{Z_k}^i$ and the knowledge of particular factor loadings of each observable individual rural development attribute (coefficient/variable) $\mathbf{VAR_a}^i$ (variables $\mathbf{VAR_a}^i$ are normally directly available or have to be computed as a coefficient (e.g., per capita) from secondary statistics on individual regions) in all $\mathbf{Z_k}$ (factorization using principal component method), the "social importance" = rank showing a *relative* contribution of each individual attribute/variable/coefficient/partial indicator ($\mathbf{R_a}^i$) to the overall rural development (at the country level) can be computed from Eq. 2:

$$\mathbf{R_a} = \Sigma_k \beta_k * \mathbf{LV_a}^k \qquad (2)$$

where

$\mathbf{R_a}$ = Relative importance (rank) of an individual regional attribute ($\mathbf{VAR_a}$) in the overall rural development (at the country level)

$\boldsymbol{\beta}_\mathbf{k}$ = <social> weight of a given factor (principal component) $\mathbf{Z_k}$ obtained from a relevant migration model

$\mathbf{LV_a}^\mathbf{k}$ = Factor loading of an individual attribute/variable/coefficient ($\mathbf{VAR_a}$) in a given factor (component) $\mathbf{Z_k}$

$\mathbf{k}$ = Number of selected factors/principal components

$\mathbf{a}$ = Index of individual attribute/variable/coefficient/partial indicator

By applying the above method, the social value of each selected partial rural development attribute $\mathbf{VAR_a}$ (i.e., contribution of individual partial indicator $\mathbf{VAR_a}$ to the overall quality of life and development level) can be measured at the country level and is equal to the weighted sum (= k) ($\boldsymbol{\beta}_\mathbf{k}$ as weights) of each attribute's respective factor loading ($\mathbf{LV_a}^\mathbf{k}$) in all selected factors/principal components $\mathbf{Z_k}$. Obviously, the combination of the *highest* factor loadings and <*highest*> social weights (in absolute terms) is decisive for the obtained rank of a given variable $\mathbf{VAR_a}$.

## Application of the RDI

The RDI developed in Michalek and Zarnekow (2012) was empirically applied to analysis of the main determinants of rural/regional development in individual rural areas in years 2002–2005 in Poland and Slovakia at NUTS-4 level.

Following this approach, the rural development domains discussed above are represented in this study by hundreds of partial socioeconomic indicators/variables (e.g., 991 variables/indicators describing various aspects of rural development at NUTS-4 level in Poland; 340 variables/indicators at NUTS-4 level in Slovakia). The constructed RDI combines all selected economic, environmental, and ▶ social indicators and links them under a consistent theoretical framework.

The data used for the calculation of the RDI for Poland originates from the Regional Data Bank of the Polish Statistical Office at (NUTS-4), as well as data obtained from the Ministry of Finance (e.g., distribution of personal income) and the Ministry of Interior (e.g., crimes) collected at NUTS-4 levels for the years 2002–2005. Of 379

NUTS-4 regions in Poland, 314 rural Powiats (NUTS-4) are included in the analysis (84.2 % of all NUTS4-regions), which excludes 65 big cities. The data basis for Poland covers all relevant rural development dimensions available in regional statistics at NUTS-4 level and consists of 991 coefficients/indicators collected/calculated either directly at NUTS-4 level or aggregated from NUTS-5 (approximately 2,500 Polish gminas) levels into NUTS-4 level.

The database for Slovakia originates from the Slovak Statistical Office whereby 337 indicators/variables collected at 72 (Okres) regions (NUTS-4) in years 2002–2005 are used for the construction of the RDI.

In both Poland and Slovakia, the number of variables characterizing various aspects of RD in individual rural regions was large and assorted regional indicators/coefficients were expected to be linearly dependent. Therefore, at the first stage the factor analysis (principles component method) was carried out. The number of retained factors in Slovakia was determined using Kaiser criterion (factors with eigenvalues greater than 1 were retained). In contrast to this procedure, the final number of selected factors in Poland was determined in an iterative procedure by selecting such a number of factors that simultaneously maximized the restricted likelihood function used in the selected model as the convergence criterion. As an outcome of factor analysis (2002–2005), 337 original variables/indicators in 72 Slovak NUTS-4 regions were converted into 21 factors characterizing various aspects (domains) of rural/regional development in Slovakia; 991 variables/coefficients in 314 rural NUTS-4 regions were converted into 17 factors in Poland. Estimated factor values in both countries are region and time specific.

An econometric estimation of weights in the RDI was carried out separately in both countries. As observable migration inflows between regions are either zero or positive, a migration function (with a dependent variable reflecting probability distribution of migration from one region to another and z-normalized factor values ($\mathbf{Z_k}$) and transaction costs modelled as a time-invariant variable consisting of two elements, i.e., distance matrix D and squared distance matrix $D^2$ – reflecting curvature properties of transaction costs – as explanatory variables) was estimated as a logistic function reflecting a probability distribution of migration from one region to another ($71 \times 72 \times 4 = 20.448$ data observations in Slovakia and $313 \times 314 \times 4 = 393.128$ data observations in Poland). The model was estimated in two versions: Version (a) as a panel regression that allows between fixed and random effect model specification (estimated as a random effects linear regression model with a group variable at the level of ID [GLS regression estimate]) and version (b) as a multilevel mixed-effect regression model (mixed-effects REML regression) that additionally allows for the possibility of the nested error structure within a region. As results on the base of models (a) and (b) for Slovakia were very similar, and model (b) is more general, final estimation results (both for Slovakia and Poland) were based on this version.

The results of this study show that among the top ten partial variables/coefficients *positively* contributing to quality of life in rural regions in Poland, the most important were:

- Personal income, i.e., the % share of the highest income group (social weight = 0.07)
- Availability and quality of new residential buildings (social weight = 0.06/0.07)
- Access to selected technical infrastructure, e.g., gas consumption from gas-line system per capita (social weight = 0.05/0.06)
- The % share (high) of the private sector in the service sector (social weight = 0.05/0.06)
- Spatial accessibility of rural enterprises (social weight = 0.05)

In Slovakia, the most important variables/coefficients *positively* contributing to local rural development were those associated with:

- Population structure (e.g., high share of population at a productive age within the total population) (social weight = 0.17/0.18)
- The % share (high) of private enterprises and natural persons in total legal units (social weight = 0.17)
- Level of consumption (high), e.g., municipal waste disposal per capita (social weight = 0.16)

- Spatial access of rural population to social infrastructure, e.g., swimming pools, sport stadia, telephone lines, post offices, and local communication, per $km^2$ (social weight = 0.1/0.12)
- The structure of local business; % share (high) of enterprises in areas: financial mediation, real estate, rental, and business activities in total enterprises (social weight = 0.12)
- Variables/coefficients associated with favorable climate and nature, e.g., high share of vineyard in agricultural land (social weight = 0.10)

Among the 10 variables/coefficients that had a particularly *negative* impact on the quality of life and rural development, the most important in Poland were:

- Low personal income: % share of low income groups (social weight = −0.07).
- The % share (high) of the public sector in the service sector (social weight = −0.06).
- Disproportion in the gender structure of the rural population, i.e., overproportional % share of male of working age (= low % share of females of working age) (social weight = −0.04).
- The % share (high) of legal units in the public administration and security sectors (social weight = −0.04).
- The % share (high) of young unemployed (25–34 years) of the total registered unemployed (social weight = −0.04).
- Level of subsidies received at gmina level (NUTS-5) (social weight = −0.03). Yet, the latter may also merely represent society's response to a low development level in the regions.

Respective variables/coefficients that were particularly *negatively* associated with the quality of life and the level of rural development in Slovakia were:

- The overproportional % share of contributory organizations and other nonprofit organizations in the structure of legal units registered in a given region (weight = −0.17). Yet, this variable (along with a number of other response variables, e.g., a high percentage of social expenditures) may also represent the policy's response to a low local development level in the past.
- The % share (high) of women among unemployed persons (weight = −0.16).
- The % share (high) of urban territory in the total area of municipality (weight = −0.13).
- The % share (high) agricultural units in total number of legal subjects registered on a given territory (weight = −0.12).
- The % share (high) of cooperatives in total enterprises (weight = −0.12).

The results of the above rankings show that the highest individual impact on the level of rural development had demographic and social domains (Poland) and the environmental and infrastructural domains (in Slovakia). On the other hand, a relatively low or even negative impact on RD was found in case of administrative variables. While economic and infrastructural domains are closely linked to each other, both of them (in total) had the highest impact on the level of rural development and the population's quality of life in rural areas in both countries. Regarding the level of regional disparities, our analysis shows that in Poland, these are very large and especially concern the best-developed regions. Indeed, the difference in estimated level of quality of life measured in terms of the RDI between the best-developed regions in Poland and a country's average was found in 2005 to be much higher than the difference in the RDI between country's average and the least-developed regions (i.e., southeast Poland).

Both groups of rural regions (i.e., the best vs. the least developed) differed significantly concerning their endowments with specific factors/principal components determining the overall quality of life. The most significant differences concerned endowments with factors: F1 (employment by sectors), F4 (highest income groups and housing availability), F6 (structure of population), F11 (primarily sector-energy, structure of deaths), F12 (population natural growth), and F16 (structure of expenditures in local budgets).

During the years 2002–2005, the estimated mean value of the RDI in Poland for 314 rural regions dropped slightly from 0.020 (2002) to

0.018 (2005) showing some fluctuation over the years. The quality of life (RDI) in the best-developed regions of rural Poland further improved (compared to the country average) while in less-developed regions deteriorated. The majority of regions which improved their absolute level of RDI were located close to bigger cities and in west and southwestern Poland (probably due to stronger socioeconomic ties with Germany and other "old" EU member states); those where the quality of life deteriorated were located mostly in northeast and eastern Poland (close to a border with Russia, Belorussia, and Ukraine) and partly in central Poland (located far from bigger cities).

The RDI constructed for Slovakia consists of 21 terms and involves 337 regional indicators calculated and weighted according to Eq. 1. During the years 2002–2005, the estimated value of the RDI ranged from –0.51 to +0.91 (regional discrepancies were therefore higher than in Poland). As expected, the highest values of RDI were found in regions located in West Slovakia (e.g., Senec, Pezinok, Dunajska Streda, Galanta), while regions of Eastern Slovakia and Central Slovakia (e.g., Gelnica, Stropkov, Namestovo, Kezmarok, Stara Lubovna) exhibited the lowest RDI values. The results also confirm a clear typographic division of Slovakia into western, central, and eastern subareas based on performance of individual regions and back up a general opinion that the level of rural development decreases from west to east. The change in the RDI across Slovak regions shows that a general pattern of development (i.e., western regions have higher RDI values compared with east Slovak regions) persisted throughout the years 2002–2005. Yet, particularly interesting was an improvement of the RDI in regions located in West and Central Slovakia, which can be interpreted as a considerable spillover effect transmitting economic and social development from better-developed regions (Western Slovakia) to less-developed regions (Central Slovakia). Yet, this encouraging development was simultaneously accompanied by an increasing variance in RDI values which indicates a progressing regional divergence.

When looking at the geographical distribution of changes in RDI by regions, our results show that most regions with an improved RDI were located in Western Slovakia and in the northern part of Central Slovakia. In Slovakia, the most significant differences between good and bad performing regions concerned endowments with factors F2 (availability of social services and technical infrastructure per capita), F3 (social and living environment including availability of housing), F10 (special schools), F4 (agriculture), F13 (public facilities), and F14 (availability of retail infrastructure). A high endowment with social and technical infrastructure calculated per capita (F2) was not found to contribute to the higher quality of life in individual rural regions (high values of regional coefficients computed per capita level may reflect a region's low population density and therefore usually do not provide reliable information about the spatial availability of a given service). Good performing regions were found to be endowed with a higher than the country's average with factors F3 (social and living environment, including availability of housing), F4 (agriculture), F13 (public facilities), and F14 (availability of retail infrastructure).

## Methodological and Policy Conclusions

The main methodological conclusions of this study are:

- An RDI allows for a comprehensive analysis of various rural development domains (economic, social, environmental, etc.) and their impact on the overall quality of life in rural regions and is powerful at NUTS-2–NUTS-5 or even village levels.
- The RDI index is not constant over time, is easily adjustable, and allows for an easy inclusion of additional relevant variables/coefficients representing various aspects of the overall quality of life/rural development.
- The weights applied into the construction of the RDI represent society's valuation of endowments and socioeconomic trends observable at local/regional levels. They are also representative for society as whole (reflects both the decision of the migrating population and of the population that stays in

the region). The weights are empirically derived and statistically verified (in the actual version the estimated weights are kept constant in time).
- The inclusion of transaction costs to the model allows for a technical separation of quality of life from migration.
- Data: an RDI is data hungry.

The main *policy* conclusion of this study is that due to its comprehensiveness and high reliability, the RDI is suitable both to an analysis of the overall level of development of rural areas as well as to a quantitative evaluation of the impacts of given RD and structural programs at regional levels. Examples of an application of the RDI for an analysis of RD programs in Poland and Slovakia at NUTS-4 level (with RDI as an impact indicator and applying matching methods, e.g., binary and generalized propensity score matching) can be found in Michalek (2012).

## References


Berger, M. C., Blomquist, G. C., & Sabirianova, K. Z. (2003). *Compensating differentials in emerging labor and housing markets*: *Estimates of quality of life in Russian cities*. Discussion Paper 900, Forschungsinstitut zur Zukunft der Arbeit, Bonn,Germany.

Douglas, S. (1997). Estimating relative standard of living in the United States using cross-migration data. *Journal of Regional Science, 37*(3), 411–436.

Douglas, S., & Wall, H. J. (1993). 'Voting with your feet' and the quality of life index: A simple non-parametric approach applied to Canada. *Economics Letters, 42* (2–3), 229–236.

Douglas, S., & Wall, H. J. (2000). The revealed cost of unemployment. *Review of the Federal Reserve Bank of St. Louis, 82*(2), 1–10.

Erikson, R. (1993). Description of inequality. The Swedish approach to welfare research. In M. C. Nussbaum & A. Sen (Eds.), *The quality of life* (pp. 67–87). Oxford, UK: Oxford University Press.

Fuguitt, G. V. (1985). The nonmetropolitan turnaround. *Annals of the Review of Sociology, 11*, 259–280.

Grasso, M., & Canova, L. (2007). *An assessment of the quality of life in the European Union based on the social indicators approach*. MPRA Paper 1785, University Library of Munich, Germany.

Greenwood, M. J. (1997). Internal migration in developed countries. In M. R. Rosenzweig & O. Stark (Eds.), *Handbook of population and family economics (chap. 12)* (pp. 647–720). Amsterdam: Elsevier Science.

Hagerty, M. R., Cummins, R. A., Ferriss, A. L., Land, K., Michalos, A. C., Peterson, M., et al. (2001). Quality of life indexes for national policy: Review and agenda for research. *Social Indicators Research, 55*(1), 1–96.

Inglehart, R. (1997). *Modernization and postmodernization: Cultural, economic and political change in 43 societies*. Princeton, NJ: Princeton University Press.

Johansson, S. (2002). Conceptualizing and measuring quality of life for national policy. *Social Indicators Research, 58*, 13–32.

Jones, A., & Riseborough, M. (2002). *Comparing quality of life research – International lessons*. Technical report, Housing Corporation.

Kaufmann, P., Stagl, S., Zawalinska, K., & Michalek, J. (2007). *Construction and the measurement of the rural development index part I: Conceptual foundations of indicators for quality of life measurement*. ADVANCED-EVAL Working Paper Series Conceptual Report CR3-1 [D 3a], University of Kiel.

Michalek J. (2012), Counterfactual impact evaluation of EU rural development programmes - Propensity Score Matching methodology applied to selected EU Member States, A regional approach, European Commission. *JRC Scientific and Policy Reports*, 2, 1–83. https://service.gmx.net/de/cgi/derefer?TYPE=3&DEST=http%3A%2F%2Fipts.jrc.ec.europa.eu%2Fpublications%2Fpub.cfm%3Fid%3D5360\t_blank http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=5360.

Michalek, J., & Zarnekow, N. (2012). Application of the rural development index to analysis of rural regions in Poland and Slovakia. *Social Indicators Research, 105*(1), 1–37.

Michalos, A. C. (2003). Migration and the quality of life: A review essay. In A. C. Michalos (Ed.), *Essays on the quality of life* (Social indicators research series, Vol. 19, pp. 89–123). Dordrecht, Germany: Kluwer Academic Publisher.

Niedomysl, T. (2006). *Migration and place attractiveness*. Ph.D. thesis, Uppsala Universitet.

OECD. (2005). *Agricultural policies in OECD countries: Monitoring and evaluation 2005. No. 2004 in agricultural policies in OECD countries*. Organisation for Economic Co-operation and Development. Paris: OECD.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy, 64*, 416–424.


## Rural Elderly

► Rural Seniors

## Rural Elders

► Rural Seniors

# Rural Life, Quality of

Antonio José López López
Políticas y Programas en Niñez, Juventud y Desarrollo Social, Universidad de Manizales-CINDE, Manizales, Colombia

## Synonyms

Individual and social perspective of rural welfare; Quality of rural living conditions, subjective perception of; Subjective quality of rural living conditions

## Definition

The quality of rural life refers to the perception of the inhabitants of rural areas – the rural – on aspects of material environment and psychosocial that they consider to be necessary to be able to move towards the gradual construction of their life projects.

## Description

The rural abandons at the end of the twentieth century its centrality in the statism and resistance to change, derived from its physical or material consideration, to be oriented as an entity of relative dimensions, modifiable, changing, derived from its union to historic and demographic expansions and contractions related to collective needs. Thus, acquiring a new centering, it considers the space/time as "recipients" or "containers" malleable and adjustable defined as concrete but transitory units of human occupancy (Fals Borda, 2000), whose activities are manifested through expressions of life influenced by its historicity.

The rural as biospace or social space and quality of life as concrete human experience are abstract representations as a result of a specific work of construction that provides a point of view on the group of points from which the ordinary agents, in ordinary behaviors, "turn their eyes to the social world" (Bourdieu, 1998: 169). In this sense, quality of rural life recognized as a sociocultural category determined by the point of view of people living in the rural, ordinary agents highly qualified to manifest through expressions that take into account the reality of the different rural processes of political, economic, or sociocultural nature (López, 2009). It focuses its attention, therefore, not only in the lives of rural people but in the perception that they posses of the way this takes place and their "capacity to achieve valuable functioning" (Sen, 1996: 56), like the construction of their life projects.

The quality of rural life has political, economic, and sociocultural connotations linked to the traditional concept of "social welfare" (Andrews & Withey, 1976; Bauer, 1966; Duncan, 1969; Michalos, 1980; Smith, 1973) and psychosocial connotations related to the evaluation that rural people do, in a moment at any given time, their positive or negative experiences, their current situations and global visions of life in general or "life satisfaction" (Diener, Suh, Lucas, & Smith, 1999: 277).

## Cross-References

- Life Satisfaction
- Quality of Life
- Social Welfare

## References

Andrews, M., & Withey, S. B. (1976). *Social indicators of well-being: Americans perceptions of life quality*. New York: Plenum Press.
Bauer, R. (1966). *Social indicators*. Cambridge, MA: MIT Press.
Bourdieu, P. (1998). *La distinción. Criterio y bases sociales del gusto*. Santafé de Bogotá: Santillana.
Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302.
Duncan, D. (1969). *Toward social reporting: Next steps*. New York: Russell Sage.
Fals Borda, O. (2000). *Acción y espacio. Autonomías en la nueva República*. Bogotá: Tercer Mundo.

López, A. (2009). Juventud rural, calidad de vida y políticas de juventud en la Zona Andina Colombiana (Versión electrónica). *Hologramática. Revista Académica de la Facultad de Ciencias Sociales de la Universidad Nacional de Lomas de Zamora, No 11, VI, pp. 61–84*. http://www.cienciared.com.ar/ra/usr/3/474/hologramatica11_v1pp61_84.pdf

Michalos, A. C. (1980). *North American social rapport: A comparative study of the quality of life in Canada and USA from 1964 to 1974*. Dordrecht: Reidel Publishing Company.

Sen, A. (1996). Capacidad y bienestar. En: Nussbaum, M. y Sen. A. (Comp.). La calidad de vida. México, D. F.: Fondo de Cultura Económica, 54–83.

Smith, D. (1973). *The geography of social well – being in the United States*. New York: McGraw-Hill.

## Rural Men and Women Well-Being in Ghana

Frank S. Arku
Development Studies, Presbyterian University College, Akuapem Campus, Akropong Akuapem, Ghana

### Definition

The well-being indicators differ among societies because they are driven by many different things including ▶ cultural values and religious and environmental conditions. Thus, well-being indicators can also be subjective in nature. It is therefore important that when determining the well-being of societies, we should not use only objective indicators but the subjective ones as well.

### Description

The international community uses standardized indicators such as literacy rates, type of energy used, access to health facilities, and income to determine people's well-being in both developed and developing countries. Nevertheless, using such standardized indicators may not always be appropriate. Applying the United Nations ▶ Human Development Index, ▶ Australia is among the top countries because of its high literacy, low mortality, and high income, but Blachflower and Oswald's (2005) study shows that most Australians' sense of ▶ well-being is compromised because of low ▶ job satisfaction. The knowledge of people's goals, objectives, and local indicators is therefore important in measuring their well-being because of the likely differences resulting from cultural values. Culture plays an important role in determining what people may consider as well-being indicators (see Shek et al., 2005; Veenhoven, 2005). The local indicators are referred to as ▶ subjective well-being indicators by Diener and Suh (2000).

Indicators of well-being can differ between urban and rural residents within a country because of differences in needs. Similarly, the indicators may differ between men and women within the same society, although there are common indicators related to the improved well-being of both men and women. While discussing the definition of ▶ development, Afshar (2005) emphasized that although material things are imperative in alleviating ▶ poverty, nonmaterial dimensions may be even more important.

Veenhoven (2005), arguing for the need of incorporation of cultural-specific indicators in determining people's ▶ quality of life, maintains that although living conditions of grain can be specified, the essential living conditions of humans are not easily determined. Accordingly, while plants need a certain amount of temperature, water, and other inputs to grow, humans can adapt when their well-being changes.

The indicators of well-being are socially and locally constructed based on the cultural values of communities (Diener & Suh, 2000; Shah, 1998; Shek et al., 2005; Veenhoven, 2005; Zorondo-Rodríguez et al., 2012). This allows for different indicators to be produced from one community to another. One hundred and fifty men and 190 women were interviewed to

**Rural Men and Women Well-Being in Ghana, Table 1** Indicators loaded high in the factor analysis

| Groups | Indicators |
|---|---|
| 1. Social | • Member of a church committee |
| | • Playing cards/drafts |
| | • Drinking palm wine together |
| | • Relaxation |
| 2. Religious and access to clean water | • Going to church regularly |
| | • Trustworthy pastor |
| | • Social interaction – they interact after church |
| | • Large family size |
| | • Close source of clean water |
| 3. Children's welfare (religious and education) | • Children going to church |
| | • Listening to radio |
| | • Children attending school |
| | • Children attending classes on time |
| 4. Economic | • Farming |
| | • Good harvests from farms |
| | • Petty trading |

Source: Arku, Filson, and Shute (2008: 378)

determine the subjective well-being indicators among rural men and women in Ghana.

## Discussion

► Factor analysis was used to reduce the well-being indicators that the respondents agreed on to the indicators that figured most prominent among their answers. The Kairser-Meyer Measure of Sampling Adequacy was used to determine whether the data were sufficient for factor analysis, and it was found that the data were adequate. This is because the result was .753 and the Bartlett's Test of Sphericity also had a very high chi-square (2272.9) which was significant at .000 level. Also, with 340 cases and 18 variables, the cases-variables ratio of approximately 19:1 is achieved.

The four main groups of well-being indicators are social, religious and access to clean water, children's welfare, and economic. While men ranked ► social indicators high, women ranked religious activities and access to clean water and children's welfare high (Table 1). Both men and women indicated that their primary well-being indicator is economic, while men mentioned farming as their primary economic well-being

**Rural Men and Women Well-Being in Ghana, Table 2** Ranking of importance of the indicators for achieving well-being between men and women

| Indicators | Gender | Mean rank | Order of rank |
|---|---|---|---|
| 1. Going to church regularly | Female | 196.0 | 3rd |
| | Male | 138.2 | 16th |
| 2. Member of a church committee | Female | 127.4 | 17th |
| | Male | 225.1 | 2nd |
| 3. Children's church attendance | Female | 166.4 | 9th |
| | Male | 175.7 | 10th |
| 4. Trustworthy pastor | Female | 196.6 | 2nd |
| | Male | 137.4 | 17th |
| 5. Farming | Female | 146.3 | 14th |
| | Male | 201.1 | 5th |
| 6. Good harvest from farms | Female | 163.0 | 12th |
| | Male | 180.0 | 7th |
| 7. Petty trading | Female | 219.8 | 1st |
| | Male | 108.1 | 18th |
| 8. Listening to radio | Female | 145.8 | 15th |
| | Male | 201.8 | 4th |
| 9. Playing cards | Female | 116.6 | 18th |
| | Male | 238.8 | 1st |
| 10. Drinking of palm wine together | Female | 130. 8 | 16th |
| | Male | 220.7 | 3rd |
| 11. Local drumming and dancing | Female | 164.5 | 10th |
| | Male | 178.1 | 9th |
| 12. Social interaction | Female | 183.3 | 5th |
| | Male | 154.3 | 14th |
| 13. Children's school attendance | Female | 151.9 | 13th |
| | Male | 194.1 | 6th |
| 14. Children going to school on time | Female | 176.5 | 7th |
| | Male | 162.9 | 12th |
| 15. Secured means of transport | Female | 163.2 | 11th |
| | Male | 179.7 | 8th |
| 16. Large family sizes | Female | 194.0 | 4th |
| | Male | 140.7 | 15th |
| 17. Close source of clean water | Female | 181.9 | 6th |
| | Male | 156.0 | 13th |
| 18. Relaxation | Female | 168.0 | 8th |
| | Male | 173.7 | 11th |

Source: Arku, Filson, and Shute (2008: 379–380)

indicator and women indicated petty trading as their primary well-being indicator. Rural Ghanaian women considered the welfare of their nuclear families as an important well-being indicator, and men considered authority as an important well-being indicator.

The Mann–Whitney test revealed that the importance of the indicators to men's and women's well-being was different (Table 2). The top five of women's indicators were petty trading, a trustworthy pastor, going to church regularly, large family sizes, and ▶ social interaction. The men's five most important indicators were playing of cards, a member of a church committee, drinking of palm wine together, listening to radio, and farming.

## References


Afshar, F. (2005). Exploring the frontiers of international development: Countries of the north, well-being, spirituality, and contemplation. *Canadian Journal of Development Studies, XXVI*, 527–545.

Arku, F. S., Filson, G. C., & Shute, J. (2008). An empirical approach to the study of well-being among rural men and women in Ghana. *Social Indicators Research, 88*, 365–387.

Blanchflower, D. G., & Oswald, A. J. (2005). Happiness and the human development index: The paradox of Australia. *The Australian Economic Review, 38*, 307–318.

Diener, E., & Suh, E. M. (2000). Measuring subjective well-being to compare the quality of life of cultures. In E. Diener & E. M. Suh (Eds.), *Culture and subjective well-being* (pp. 3–12). London: The MIT Press.

Shah, M. K. (1998). Gendered perceptions of well-being and social change in Darko, Ghana. In I. Guijt & M. K. Shah (Eds.), *The myth of community: Gender issues in participatory development* (pp. 141–151). London: Intermediate Technology Publications.

Shek, T. L., Chan, Y. K., & Lee, P. S. N. (2005). Quality of life in the global context: A Chinese response. *Social Indicators Research, 71*, 1–10.

Veenhoven, R. (2005). Apparent quality-of-life in nations: How long and happy people live. *Social Indicators Research, 71*, 61–86.

Zorondo-Rodríguez, F., Gómez-Baggethun, E., Demps, K., Arìza-Montobbio, P., García, C., & Reyes-García, V. (2012). What defines quality of life? The gap between public policies and locally defined indicators among residents of Kodagu, Karnataka (India). *Social Indicators Research* (Online)


# Rural Seniors


Bonnie Jeffery[1], Juanita Bacsu[2], Sylvia Abonyi[3], Shanthi Johnson[4], Diane Martz[5] and Nuelle Novik[6]

[1]Faculty of Social Work, University of Regina, Prince Albert, SK, Canada

[2]Saskatchewan Population Health and Evaluation Research Unit, University of Saskatchewan, Saskatoon, SK, Canada

[3]University of Saskatchewan, Saskatoon, SK, Canada

[4]Faculty of Kinesiology and Health Studies, University of Regina, Regina, SK, Canada

[5]University of Saskatchewan, Saskatoon, SK, Canada

[6]University of Regina, Regina, SK, Canada


## Synonyms

Nonurban seniors; Older adults in rural areas; Older people in rural areas; Rural elderly; Rural elders; Seniors in remote areas

## Definition

The phrase "rural seniors" refers to older adults (age 65 and older) who live in rural communities which are underserviced, sparsely populated, geographically disperse, and distant from large urban centers. Rural seniors report challenges related to access to healthcare, housing, public transportation, caregiver supports and information on existing services (Bacsu et al., 2012). Multiple levels of determinants (kin/family, community, and policy) influence rural seniors' quality of life and need to be examined in order to support healthy aging in place. Research on rural aging needs to address rural seniors' diversity based on individual and community characteristics and Aboriginal and non-Aboriginal status.

## Description

Seniors are one of the fastest growing population groups in Canada which makes understanding their quality of life and health needs of vital importance. Quality of life is often defined as a multidimensional concept that relates to an individual's life satisfaction and overall well-being (Ory, Cox, Gift, & Abels, 1994). The term rural is often used to describe communities located outside of urban areas which are underserviced, sparsely populated, and geographically dispersed (Kivett, Stevenson, & Zwane, 2000). A study by Statistics Canada projects that between 2005 and 2036, the number of seniors in Canada will increase from 4.2 million to 9.8 million (Statistics Canada, 2006, p. 12). The study also found that rural regions were aging the fastest in terms of their growth in the number of seniors (Statistics Canada, 2006, p. 10).

An important concept that relates to ▶ quality of life for rural seniors is "healthy aging in place" (Shulamit, Zimmerman, & Eckert, 2001). This describes the situation where seniors are provided with the necessary home and community supports so that they have the choice and ability to remain in their communities as long as they wish. There is, however, a growing uncertainty around the ability of rural areas to support and maintain healthy aging in place (Eberhard & Pamuk, 2004) which highlights the importance of research that can contribute to a greater understanding of these health needs from the perspective of seniors and that can synthesize best practices in terms of policies and programs to meet these needs.

The quality of life experienced by rural seniors in Canada is influenced by individual- and community-level determinants. As the Canadian rural population ages, there is a significant need for policy-makers to develop programs and policies that improve the quality of life for rural seniors. Aspects such as individual health needs, availability of and access to health services, availability of support systems, and availability of appropriate and affordable housing are key elements that influence quality of life for rural seniors.

### Health Needs

Studies have reported that rural residents have shorter life expectancies, higher death rates, and higher infant mortality rates than residents of urban areas (DesMeules et al., 2006). Rural areas also experience challenges in providing sufficient health service provision due to the physical and remote geography, poor infrastructure, hostile terrain, limited tax base, and small and dispersed populations (Bull, Krout, Rathbone-McCuan, & Shreffleev, 2001). Accordingly, many rural communities in Canada are facing significant threats in meeting the health needs of an aging population.

Rural seniors are not a homogenous group, and we find that there are differences between non-Aboriginal rural senior women and men on various health outcome measures. Recent studies show that in comparison to urban seniors and rural senior women, rural senior men are often severely disadvantaged in terms of having shorter life expectancies (DesMeules et al., 2006), higher suicide rates (Eberhard & Pamuk, 2004), higher injury rates (Amshoff & Reed, 2005), and participate in fewer health screenings and allow illness to progress longer (Denner, 2009). Less research has been conducted among Aboriginal seniors, but the general relationship between poor health and age is similar to non-Aboriginal seniors, with older Aboriginal peoples in general having poorer health status and suffering up to triple the rate of chronic health conditions such as hypertension, heart problems, and diabetes than their non-Aboriginal counterparts over all (Rosenberg et al., 2008).

### Health Services

Studies suggest that limited access to medical services in rural areas may be an involuntary barrier to rural seniors' usage of medical services and preventative screening measures. People living in rural areas usually travel long distances to receive health services or have to wait to see medical professionals who may occasionally visit the area. Wong and Regan (2009) suggest that rural people needing additional health services to manage their health problems often make tradeoffs between their own safety in travelling during times of poor road conditions and having their health-care needs met.

### Support Systems

Predictions of escalating pressures on rural health services continue to rise with the anticipated growth of seniors in Canada. Support systems and community-level supports are increasingly being recognized as an essential component in understanding older adults' health and well-being (Ashida, 2008). Isolation and lack of support from others has long been connected with depression, heart disease, health-risk behaviors, and a variety of other health outcomes (Minkler & Wallerstein, 2004). To date, little research exists on community-level interventions that support healthy aging in place. Kirby and LeBreton (2002, p. 151) note that often "rural health issues tend to be eclipsed by those in urban areas... policy solutions often are based on experiences in urban areas and rely on urban data and research." Nordberg (2007), for example, reports that rural seniors rely more on informal, community-level interactions than their urban counterparts.

Cantor's (1989) Social Care Model is a useful framework for understanding the different levels of networks within a social support system for rural seniors. First, formal networks are conceptualized as governmental organizations, voluntary services, political institutions, and economic organizations. Second, mediating support elements are defined as the quasi-formal networks and "include religious organizations, racial/cultural, social, neighborhood, and block groups as well as individuals such as mail carriers, shop keepers, bartenders and building superintendents" (1989, p. 103). Third, informal networks are described as kin, friends, and neighbors. In this model, the "older persons interact with each of these circles at varying times and for varying types of assistance, and at times these separate networks interact and even overlap" (1989, p. 103).

Currently most studies have conceptualized the concept of seniors' support networks to be homogenous (Kivett et al., 2000; Statistics Canada 2002). However, there is significant evidence that there are important differences among social, emotional support and personal care networks that need to be accounted for (Dobbs, Swindle, Keating, Eales, & Keefe, 2004). For Aboriginal seniors, the importance of retaining aging seniors in communities and supporting their healthy aging in place produces differences in the nature and value of informal care systems (community and kin based) and in the timing and role of structured and formal care (e.g., institutionalized, home care) compared to non-Aboriginal seniors (Buchignani & Armstrong-Esther, 1999).

### Appropriate and Affordable Housing

Many rural communities are experiencing facing challenges in meeting the needs of appropriate housing for seniors. The demand for rural seniors' housing is stimulated by seniors' desire to age in their communities, in-migration of seniors retiring from urban communities, farmers moving into town, and people moving back to the towns they were raised in (Statistics Canada, 2006, p. 10). In comparison to urban seniors, most rural seniors own their own homes; however, their homes are often older and in need of costly repair (Bull et al., 2001). Rural seniors looking to move out of their homes may experience challenges in finding housing that is safe, affordable, and accessible.

Rural communities in Canada face escalating challenges in meeting the health-care needs of an aging population. There is a growing uncertainty around the ability of rural areas to support and maintain seniors' health-care needs and quality of life (Eberhard & Pamuk, 2004). Despite this situation, there are few studies that directly address rural seniors' health, and there are significant knowledge gaps in our understanding of non-Aboriginal rural seniors' quality of life. We know even less about the situation of Aboriginal seniors where, for example, Rosenberg et al. (2008) report that little has been written about older Aboriginal peoples where older Aboriginal peoples are the explicit focus of the research – conceptually or materially in any context, such as urban, rural, or northern. Accordingly, there is a need for research that can contribute to a greater understanding of non-Aboriginal and Aboriginal rural seniors' health needs and their quality of life.

## Cross-References

- ▶ Aging Population
- ▶ Community Support
- ▶ Elderly Activity and Engagement with Life
- ▶ Health Determinants
- ▶ Health Inequities
- ▶ Housing and Aging
- ▶ Indigenous Health Disparities
- ▶ Social Support

## References


Amshoff, S., & Reed, D. (2005). Health, work, and safety of farmers ages 50 and older. *Geriatric Nursing, 26*(5), 304–308.

Ashida, S. (2008). Differential associations of social support and social connectedness with structural features of social networks and the health status of older adults. *Journal of Aging and Health, 20*(7), 872–893.

Bacsu, J., Jeffery, B., Johnson, S., Martz, D., Novik, N., & Abonyi, S. (2012). Healthy aging in place: Supporting rural seniors' health needs. *Journal of Rural Nursing and Healthcare, 12*(2), 77–87.

Buchignani, N., & Armstrong-Esther, C. (1999). Informal care and older native Canadians. *Ageing and Society, 19*, 3–32.

Bull, N., Krout, J., Rathbone-McCuan, E., & Shreffleev, M. (2001). Access and issues of equity in remote/rural areas. *The Journal of Rural Health, 17*(4), 356–359.

Cantor, M. H. (1989). Social care: Family and community support systems. *The Annals of the American Academy of Political and Social Science, 503*, 99–112.

Denner, B. (2009). *Reducing early mortality of men living in rural and remote Australian communities*. 10th National Rural Health Conference. Retrieved September 13, 2011, from http://10thnrhc.ruralhealth.org.au/papers/docs/Denner_Bernard_D7.pdf

DesMeules, M., Pong, R., Lagacé, C., Heng, D., Manuel, D., Pitblado, R., et al. (2006). *How healthy are rural Canadians? An assessment of their health status and health determinants*. Ottawa, ON: Canadian Institute for Health Information.

Dobbs, B., Swindle, J., Keating, N., Eales, J., & Keefe, J. (2004). *Caring context of rural seniors – Phase II technical report*. Veterans Affairs Canada. Edmonton: University of Alberta.

Eberhard, M., & Pamuk, E. (2004). The importance of place of residence: Examining health in rural and nonrural areas. *American Journal of Public Health, 94*(10), 1682–1686.

Kirby, M. & LeBreton, M. (2002). *The health of Canadians – The federal role. Volume two: Current trends and future challenges*. Ottawa, ON: Standing Senate Committee on Social Affairs, Science and Technology. Retrieved September 10, 2011, from http://www.parl.gc.ca/37/1/parlbus/commbus/senate/com-e/SOCI-E/rep-e/repjan01vol2-e.pdf

Kivett, V., Stevenson, M., & Zwane, C. (2000). Very-old rural adults: Functional status and social support. *Journal of Applied Gerontology, 19*(1), 58–77.

Minkler, M., & Wallerstein, N. (2004). Improving health through community organizations and community building: A health education perspective. In N. Minkler (Ed.), *Community organizations and community building for health* (2nd ed., pp. 30–52). New Brunswick, NJ: Rutgers University Press.

Nordberg, G. (2007). *Formal and informal care in an urban and rural elderly population: Who? When? What?* Stockholm: Aging Research Center. Retrieved September 10, 2011, from http://www.aldrecentrum.se/upload/Nyheter/L%C3%A4nkar%20til/Avh_formal_informal_care.pdf

Ory, M., Cox, D., Gift, H., & Abels, R. (1994). Aging and quality of life: Celebrating new discoveries. In R. Abeles & H. Gift (Eds.), *Aging and quality of life* (pp. 1–16). New York: Springer.

Rosenberg, M., Wilson, K., Abonyi, S., Wiebe, A., Beach, K., & Lovelace, R. (2008). *Older aboriginal peoples in Canada – Demographics, health status and access to health care* [SEDAP Research Paper No. 249]. pp. 1–24. Retrieved September 13, 2011, from http://socserv.mcmaster.ca/sedap/papers09.htm

Shulamit, L. B., Zimmerman, S., & Eckert, J. K. (2001). Aging in place. In S. Zimmerman, P. D. Sloane, & J. K. Eckert (Eds.), *Assisted living needs, practices and policies in residential care for the elderly*. Baltimore: The Johns Hopkins University Press.

Statistics Canada. (2002). *General Social Survey Cycle 16: Caring for an aging society*, Available online Retrieved April 6, 2013, from http://prod.library.utoronto.ca/datalib/codebooks/cstdli/gss/gss16/89-582-XIE03001.pdf.

Statistics Canada. (2006). *A portrait of seniors in Canada*. Retrieved from September 10, 2011, http://www.statcan.gc.ca/pub/89-519-x/89-519-x2006001-eng.pdf

Wong, S. & Regan, S. (2009). Patient perspective on primary health care in rural communities: Effects of geography on access, continuity and efficiency. *Rural and Remote Health, 9*, 1–12. Retrieved September 13, 2011, from http://www.rrh.org.au


# Russian America

- ▶ Alaska, Living Conditions of the Inupiat