

SOEPpapers

on Multidisciplinary Panel Data Research

145

Ulrich Schimmack

Measuring Wellbeing in the SOEP

Berlin, November 2008

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
<http://www.diw.de/soeppapers>

Editors:

Georg **Meran** (Vice President DIW Berlin)

Gert G. **Wagner** (Social Sciences)

Joachim R. **Frick** (Empirical Economics)

Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics)

Christoph **Breuer** (Sport Science, DIW Research Professor)

Anita I. **Drever** (Geography)

Elke **Holst** (Gender Studies)

Frieder R. **Lang** (Psychology, DIW Research Professor)

Jörg-Peter **Schräpler** (Survey Methodology)

C. Katharina **Spieß** (Educational Science)

Martin **Spieß** (Survey Methodology)

Alan S. **Zuckerman** (Political Science, DIW Research Professor)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | urahmann@diw.de

Measuring Wellbeing in the SOEP

By Ulrich Schimmack*

University of Toronto Mississauga

Abstract

I define wellbeing as preference realization. Wellbeing can be measured with affective (the amount of pleasant versus unpleasant experiences) and cognitive (satisfaction with life in general and life domains) measures. Since its inception 25 years ago, the SOEP has included cognitive measures of wellbeing. In 2007, the SOEP included four items (happy, sad, angry, afraid) as an affective measure of wellbeing. This paper examines similarities and differences between cognitive and affective measures of wellbeing. In the end, I propose a wellbeing index that combines information from measures of life satisfaction, average domain satisfaction, and affect balance.

JEL Classification: I31 - General Welfare; Quality of Life; Happiness; Wellbeing

* The data used in this publication were made available to us by the German Socio- Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW Berlin).

Introduction

The concept of wellbeing has deep roots in philosophy. Modern definitions of wellbeing emerged in the 19th century. The main contribution of the utilitarian movement was to define wellbeing subjectively and to proclaim individuals' wellbeing as an important, if not the only, goal of individuals' behavior and public policy. During the 20th century social scientists started to examine wellbeing empirically, but a unified concept of wellbeing was lacking. At the beginning of the 20th century, economists developed elaborate quantitative theories of wellbeing (utility), but rejected the possibility that individuals' could provide valid reports of their own wellbeing. In the second half of the 20th century social scientists started to develop subjective measures of wellbeing, and started to examine how these measures relate to demographic variables or other characteristics of individuals (Andrews & Withey, 1976; Cantril, 1965).

Over the past decades, a large body of empirical evidence on correlates of wellbeing measures has accumulated (Diener, Suh, Lucas, & Smith, 1999). Unfortunately, the empirical evidence does not provide a solid foundation for the development of theories or policy recommendations. One major reason is the lack of a clear definition of wellbeing and insufficient research on the validity of wellbeing measures. Validation research requires a clear definition of a construct. For example, a scale is a valid measure of weight, but not a valid measure of intelligence. Thus, it is important to specify the construct of weight to examine the validity of a scale as a measure of weight. Similarly, it is impossible to examine the validity of wellbeing measures without defining wellbeing.

What is Wellbeing?

It is impossible to review the major theories of wellbeing here. Sumner (1996) provides a good overview and classification of various wellbeing definitions. His first distinction is between objective and subjective definitions of wellbeing. The distinction is based on the selection process of the criteria that are used to judge individuals' wellbeing. Objective definitions assume that the criteria can be defined without reference to the individual's own preferences, interests, ideals, values, and attitudes. Subjective definitions require that individuals' preferences, interests, ideals, values, and attitudes matter. Without going into details here, I agree with various other wellbeing scientists as well as philosophers that wellbeing has to be defined subjectively because objective definitions encounter insurmountable problems (Diener, 1984; Schimmack, 2008, Sumner, 1996).

The most important distinction among subjective theories is whether they focus exclusively on mental states or also incorporate actual states of the world in the concept of wellbeing. This criterion separates traditional definitions of utility in terms of pleasure and pain (Kahneman, 1999) from preference realization (Schimmack, 2008). What is at stake here is how wellbeing science should deal with feelings and other mental states that are based on illusory perceptions of reality. Some psychologists have argued that positive illusions are common, normal, and healthy (Taylor & Brown, 1988). If good feelings were the only criterion for high wellbeing, then promoting good feelings via illusory beliefs would be one strategy to promote wellbeing.

Other theorists have argued that mental states may be insufficient to define wellbeing (Sumner, 1996). To use Sumner's (1996) example, a man may feel happy because he assumes that his wife loves him and is faithful to him, when this is actually not the case. However, many people may prefer knowing about the infidelity and feeling unhappy about it over illusory happiness. In this case, the positive feelings based on false perceptions of reality would not be an indicator of wellbeing, if wellbeing is defined as the actual realization of subjective preferences (Schimmack, 2008). Preference realization can include mental states and feelings because people can have preferences for some feelings over others. Most people prefer pleasure to pain most of the time. It can also allow for illusions to have positive effects on wellbeing if people have a preference to be ignorant and happy rather than informed and unhappy about things that are objectively not going well in their lives.

Measuring Wellbeing

In the previous section, I argued that wellbeing cannot be reduced to the amount of pleasant and unpleasant feelings that people actually experience. There are two reasons why hedonic measures of wellbeing are still useful, even if wellbeing is defined as preference realization (Schimmack, 2008). One reason is that many people have a strong preference to feel good. Empirical studies suggest that this preference has also gained in importance over time and with increasing wealth (Inglehart, 1997). However, this argument does not give hedonic feelings a special status. It could be equally important to measure other aspects of human's life that could reflect preference-realization (e.g., health).

A more powerful argument to assess hedonic feelings is that people's feelings respond to their life circumstances (Lazarus, 1991). Moreover, the affective response depends on people's preferences. A car driving by blasting hip-hop music could produce feelings of displeasure and irritation for one person and feelings of pleasure and enjoyment for another person. Thus, people's feelings are one indicator of their preference realization.

Hedonic measures are not the only way to measure wellbeing. A more direct way to assess preference realization is to directly ask people to evaluate their lives based on their personal preferences (Cantril, 1965), or to ask for satisfaction with various life domains that are important to most people (Andrews & Whithey, 1976). Finally, it is possible to assess wellbeing by assessing people's opportunities to realize their preferences either in terms of their monetary resources or more broadly by also taking education, health, and freedom into account.

Diener, Lucas, Schimmack, and Helliwell (2008) point out the various strengths and weaknesses of different measures. Most importantly, none of these measures can claim a priori to be the ultimate or even the best measure of wellbeing. Thus, an important question is how well various measures of wellbeing actually measure wellbeing. Ample evidence demonstrates the convergent validity of various wellbeing measures. Convergent validity is reassuring, especially if two different methods are used. However, convergent validity is not always high, and when self-ratings are used, estimates of convergent validity can be biased by shared method variance. Many wellbeing measures also show convergent validity with ratings by informants, but these correlations can be quite modest, often not exceeding .5 (Schneider & Schimmack, 2008). Thus, the empirical evidence shows that wellbeing

measures have more validity than behaviorists and other skeptics considered possible, but the evidence also leaves ample room for the influence of distortions and biases.

Confusion of Measures and Constructs

Unfortunately, wellbeing scientists who use subjective measures have often ignored the fact that their measures are only imperfect indicators of wellbeing. Often the measure is equated with the construct, which is a common fallacy in the social sciences (Borsboom, 2006). Ignoring measurement problems can produce misleading results. For example, rankings of nations in the world value survey vary quite dramatically depending on the choice of the wellbeing measure. Despite high reliability of national averages, a 10-point life-satisfaction scale and a four-point happiness scale produce different results (Deaton, 2008; Inglehart, Foa, Peterson, & Welzel, 2008). Thus, it becomes important to determine the factors that produce discrepancies between different measures of wellbeing.

A Comparison of Multiple Measures of Wellbeing

This paper examines the validity of wellbeing measures by examining convergence and discrepancies of several well-being measures. If a diverse set of measures produce similar findings, it is more likely that the results would also generalize to the unobserved variation in wellbeing. However, if different measures produce different results, a careful examination of the discrepancies is required.

Global life-satisfaction. The first measure is the global 11-point rating of life-satisfaction (e.g., Schimmack, Schupp, & Wagner, 2008). This item is nearly exclusively used as a measure of wellbeing in the SOEP, and it is sometimes assumed that this measure can be treated as a cardinal measure of utility. The reasons for its popularity are its high face validity and the widespread use of life-satisfaction ratings in the wellbeing literature. Moreover, life-satisfaction ratings have shown impressive validity, and many potential biases have been shown to have relatively little effect on these ratings (Schimmack & Oishi, 2005). However, this evidence does not suggest that life-satisfaction judgments are the best measures of wellbeing, and some studies have revealed systematic biases in life-satisfaction judgments (Schimmack et al., 2008).

Affect balance. The second measure is an affective measure of wellbeing. Theoretical considerations and pilot testing produced an affect balance measure based on four items. One item assesses positive experiences (happy), whereas the other three items assess negative experiences (angry, afraid, sad). The pilot study also suggested that subtracting the average of the three negative items from the positive item produced a good measure of affect balance. The measure is highly correlated with a longer 10-item measure and produces results that are consistent with findings in the psychological literature (Schimmack et al., 2008).

Average domain satisfaction. The third measure is an average of various domain satisfactions that are routinely assessed in the SOEP (health, work, household income, leisure time, family). Previous studies have shown high correlations between global life-satisfaction judgments and judgments of average domain satisfaction (Andrews & Whitey, 1976; Schimmack, Diener, & Oishi, 2002; Schimmack & Oishi, 2005). This measure has two drawbacks. First, it does not weigh domains by the subjective importance of domains. Second, the measure fails to capture aspects of wellbeing that are not covered

by the domains included in the survey. A main advantage of this measure is that it relies less on respondents' ability to summarize and weigh all relevant aspects of their lives in response to a single question about satisfaction with life in general. Moreover, unweighted averages can be surprisingly robust estimates of weighted averages (Andrews & Whithey, 1976; Schimmack et al., 2002).

Income. The fourth measure is income. Although psychologists do not regard income as a measure of wellbeing, income is an important indicator in (welfare) economics. Briefly, income is considered an important factor that influences wellbeing because it enhances people's opportunity to realize their preferences. However, standard economists are aware that income is an imperfect indicator of wellbeing for a number of reasons. One important reason is that not all preferences can be realized with money for ethical (e.g., it is illegal to pay a contract killer) or logical (it is impossible to buy a loyal friend or unconditional love) reasons. The second reason is that salaries could be negatively related to other aspects of a job that influence preference realization. Finally, income is an imperfect measure of spending and consumption, which are the more proximal predictors of wellbeing. Although the relation between income and wellbeing is not perfect, a wide variety of preferences can be realized better with more money (a better house, a shorter commute, fancier vacations, better schools for one's children, more leisure by paying for help with chores, etc.). Not surprisingly, household income is consistently positively related with global happiness and life-satisfaction ratings (Diener et al., 1999). Across nations, income measured in GDP is a very strong predictor of average life-satisfaction ratings (Deaton, 2008).

Method

The results reported in this article are based on the 2006 and 2007 waves of the SOEP. The analyses are based on respondents with complete data on the variables of interest. Detailed documentation of the variables can be found on the DIW website (SEOP Info). Data from 2006 and 2007 were used to examine the retest reliability of measures that were administered in both years. This is especially important for global life-satisfaction because this single item measure has a relatively low reliability of about .6 (Schimmack, Wagner, Krause, & Schupp, 2007).

Results

Table 1
Correlations among wellbeing measures in 2006 and 2007 ($N = 18,196$)

Measure (Year)	LS06	LS07	DS06	DS07	AB07	INC06	INC07
LS06 Global Life-Satisfaction (2006)	-						
LS07 Global Life-Satisfaction (2007)	.60	-					
DS06 Average domain satisfaction (2006)	.67	.55	-				
DS07 Average domain satisfaction (2007)	.56	.68	.70	-			
AB07 Affective Balance (2007)	.42	.54	.43	.53	-		
INC06 Household Income (2006)	.25	.24	.24	.23	.14	-	
INC07 Household Income (2007)	.25	.25	.23	.24	.16	.86	-

Table 1 shows the correlations among the various measures. The first important finding is the range of convergent validity coefficients from $r = .68$ for global life-satisfaction and average domain satisfaction in 2008 to $r = .14$ for household income in 2006 and affect balance in 2007. The second noteworthy finding is that correlations among self-report measures are influenced by the year of assessment, although the effect size of the difference is relatively small ($Q \sim .10$). In contrast, correlations with income are essentially identical across years ($Q < .03$). The nearly identical correlations with income are due to the high temporal stability of income. The finding that subjective reports systematically change in the presence of stable economic conditions can be interpreted in two ways. It could reveal that self-report measures are sensitive to the influence of changes in non-monetary influences on wellbeing, but it is also possible that shared method variance among self-report measures contributes to the results. The third finding is that cognitive measures are more strongly correlated with income than affect balance, although the difference is again small ($Q \sim .10$). This finding also can be interpreted in two ways. One interpretation is that cognitive measures are biased because satisfaction with income is weighted too heavily (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2006). The other interpretation is that affect balance is only a partial indicator of preference realization. Moreover, the shorter time frame of the measure makes it more sensitive to transitory factors that influence wellbeing. Finally, the results show that global life-satisfaction and average domain satisfaction are approximately equally highly correlated with affect balance and income ($Q < .03$). This finding suggests that both measures are equally valid measures of wellbeing.

Correlations with Other Variables

I created a composite subjective measure of wellbeing by combining average domain satisfaction (.3), life-satisfaction (.3), and affect balance (.4). The weights weigh cognitive indicators slightly more than the affect balance measure, but results would be quite similar if both components were weighted equally. Table 2 shows how the various wellbeing indicators correlate with variables that have been related to wellbeing in other studies. The most important finding is that strong predictors (e.g., unemployment) produce consistent correlations with all

Table 2

Correlations of wellbeing measures in 2007 with predictor variables (N = 15,363).

Predictor	DS	LS	AB	SWB
Neuroticism	-.23	-.24	-.32	-.31
Unemployment	-.16	-.18	-.10	-.16
Disability	-.16	-.18	-.14	-.18
Divorce	-.12	-.09	-.08	-.11
Widowhood	-.03	-.04	-.09	-.07
Gender (0=M/1=F)	.00	.01	-.09	-.04
Birth Year	.01	.08	.08	.07
Income	.24	.26	.17	.25

Note. Average domain satisfaction (DS), life satisfaction judgment (LS), affect balance (AB), subjective wellbeing measures (SWB).

measures although the effect size varies slightly. In contrast, weak predictors (e.g., gender) produce inconsistent results. The main implication is that it is impossible to make empirically supported claims about the relation of these variables with wellbeing because it is unclear which measure produces valid results. Another noteworthy finding is that some predictors are stronger predictors of cognitive indicators (unemployment), whereas others are stronger predictors of affective indicators (neuroticism, widowhood). These patterns can be exploited to test causal theories of wellbeing, but they also create problems for the measurement of wellbeing (Schimmack et al., 2008). In the absence of more precise measures, the SWB index is likely to provide the most valid results because it uses the strength of various approaches to measure wellbeing. Future research should continue to examine the validity of its components and create more precise measures of wellbeing.

References

- Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being*. New York: Plenum.
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, 71(3), 425-440.
- Cantril, H. (1965). *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press.
- Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the Gallup World Poll. *Journal of Economic Perspectives*, 22(2), 53-72.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575.
- Diener, E., Lucas, R. E., Schimmack, U., & Helliwell, J. (2008). *Well being and public policy*. Oxford University Press. (in press).

- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*(2), 276-302.
- Inglehart, R. (1997). *Modernization and post-modernization*. Princeton, NJ: Princeton University Press.
- Inglehart, R., Foa, R., Peterson, C., & Welzel, C. (2008). Development, freedom, and rising happiness. *Perspectives on Psychological Science*, *3*, 264-285.
- Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*. (pp. 3-25): Russell Sage Foundation, New York, NY, US.
- Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. A. (2006). Would you be happier if you were richer? A focusing illusion. *Science*, *312*, 1908-1910.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Schimmack, U. (2008). Wellbeingscience.org - The science of wellbeing.
- Schimmack, U., Diener, E., & Oishi, S. (2002). Life-satisfaction is a momentary judgment and a stable personality characteristic: The use of chronically accessible and stable sources. *Journal of Personality*, *70*(3), 345-384.
- Schimmack, U., Krause, P., Wagner, G. G., & Schupp, J. (2007). *Stability and change of SWB: An experimentally enhanced latent state-trait-error analysis*. Submitted for Publication.
- Schimmack, U., & Oishi, S. (2005). The influence of chronically and temporarily accessible information on life satisfaction judgments. *Journal of Personality and Social Psychology*, *89*(3), 395-406.
- Schimmack, U., Schupp, J., & Wagner, G. G. (2008). The influence of environment and personality on the affective and cognitive component of subjective well-being. *Social Indicators Research*, *89*, 41-60.
- Schneider, L., & Schimmack, U. (2008). *Self-informant agreement in well-being ratings: A meta-analysis*. Submitted for Publication.
- Sumner, L. W. (1996). *Welfare, happiness, and ethics*. Oxford: Clarendon Press.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193-210.