



RatSWD

Working Paper Series

Working Paper

No. 119

Subjective Indicators

Beatrice Rammstedt

August 2009

Working Paper Series of the Council for Social and Economic Data (RatSWD)

The *RatSWD Working Papers* series was launched at the end of 2007. Since 2009, the series has been publishing exclusively conceptual and historical works dealing with the organization of the German statistical infrastructure and research infrastructure in the social, behavioral, and economic sciences. Papers that have appeared in the series deal primarily with the organization of Germany's official statistical system, government agency research, and academic research infrastructure, as well as directly with the work of the RatSWD. Papers addressing the aforementioned topics in other countries as well as supranational aspects are particularly welcome.

RatSWD Working Papers are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The *RatSWD Working Papers* are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the *RatSWD Working Papers* website presents only those papers published in English, while the German section lists the complete contents of all issues in the series in chronological order.

Starting in 2009, some of the empirical research papers that originally appeared in the *RatSWD Working Papers* series will be published in the series *RatSWD Research Notes*.

The views expressed in the *RatSWD Working Papers* are exclusively the opinions of their authors and not those of the RatSWD.

The RatSWD Working Paper Series is edited by:

Chair of the RatSWD (2007/ 2008 Heike Solga; 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)

Subjective Indicators

Beatrice Rammstedt

GESIS – Leibniz-Institut für Sozialwissenschaften (beatrice.rammstedt[at]gesis.org)

Abstract

Subjective indicators have been proven to possess predictive power for a large array of social and economic outcomes. However, most of these measures face serious psychometric shortcomings, namely that the items used are not psychometrically investigated. Further, for the assessment of one and the same construct various different item phrasing and response formats are used in different surveys. In the present paper several recommendations are made to increase the quality and by that also the acceptance and usage of subjective indicators. These include to develop more ultra-short but multi-item measures for subjective indicators. Further, surveys should try to use the same form of measurement, i.e. the same item phrasings and the same response scales. In terms of psychometric properties it is recommended to investigate reliability and validity of the indicators in as much depth as possible. In addition suggestions are made how to investigate the respondent's judgmental process for the subjective indicators' measures which allows to obtain a clearer picture of how the item is understood by the respondent and on which cues he bases his judgment.

Keywords: subjective indicators, reliability, validity, multi-item instruments, cognitive interviews

Certainly one can assess someone's state of health by investigating his patient records, the number of sick days or the amount of pharmaceuticals the person takes. These, however, are only proxies for the person's state of health. Yet, what is more important is the perceived, thus the subjective health. A person going to work even though he feels sick, who is not taking medication and does not see a doctor will be much less productive than a person who perceives himself as healthy.

Definition: What are subjective indicators?

At first sight it might be assumed that the distinction between objective and subjective indicators is rather unambiguous. Objective measurement is based on explicit criteria and is performed by external observers (cf. Veenhofen 2007). An example could be again the state of health, assessed by objective indicators such as antigens in the blood or the income level, assessed by the annual pay check. However, such objective criteria may also be assessed by subjective measures, e.g. by use of self-ratings. Are such self-ratings and thus subjective measures of health or income still objective or do they become subjective measures? The other way round one could think of subjective criteria, such as the individually perceived insecurity assessed by objective measures, such as the installation of alarm systems or the training in self-defense. Would this be an objective or a subjective indicator? We thus need to differentiate between objective and subjective on both, the side of the criterion or substance to be assessed and the assessment itself. Both can be either objective or subjective.

Subjective indicators are often defined as information that includes some kind of a subjective component, such as a personal perception or a personal evaluation (cf. Noll 2001). This definition is thus only focused on the subjective measurement of a whatsoever natured criterion. Following the above described differentiation, this definition includes subjective measures of both, objective as well as subjective criteria. For recent understandings this definition is far too broad as it would include any self-report data and could thus be in no way defined by the criterion to be rated. In the following, subjective indicators are defined as subjective measures of a subjective criterion. Thus, a comparably strong definition is chosen which however enables to narrow the broad and often somewhat vague field of subjective indicators. Subjective indicators – in this sense – are thus defined as *subjective information about a subjective criterion*. Subjective indicators thus include constructs like satisfaction, worries or trust.

The reason for assessing subjective indicators

Why do we assess subjective indicators at all, although we know that subjective measures like self-ratings have strong shortcomings as they are – as the labeling says – subjective. To answer this question we need to disentangle the subjective indicator again in the subjective substance and its subjective measurement.

Why are researchers interested in subjective substances such as happiness, worries or values? It is commonly objected that such matters are unstable over time, are incomparable across cultures and even individuals, and are unintelligible, as they are implicit and thus their ‘true score’ can hardly be investigated (cf. Veenhoven 2007). However, research findings have shown that these objections do not apply to *all* subjective criteria. There seem to be several subjective constructs, e.g. overall life satisfaction, of which individuals have a clear, stable and comparable understanding. So that even though the rating itself is subjective the construct it is based on seems to be comparable across individuals. Furthermore, these subjective criteria often have a high impact on external, objective criteria such as objective well-fare or suicide rate. Thus, for the prediction of several important life or societal outcomes, the measurement of subjective substances is indispensable.

However, the question might be raised, why these subjective substances are measured by self-reports and not by means of more objective data, such as an indicator for the quality of living conditions to estimate well-being. One reason is, that it is often much more time-consuming to gather information on such objective measures than simply to ask the respondent about his/her well-being. The most important reason, however, is that such objective indices do not seem to be an appropriate proxy for the individually perceived criterion. This individual’s ‘true score’ in the subjective criterion is only partially based on objective measures as, for example, life circumstances. The other part of this ‘true score’ is based on several aspects such as the individual’s fitness to its environment and to its capacity to cope with these life circumstances. This latter part, however, can hardly be assessed by objective proxies.

In sum, research has shown that subjective criteria are fruitful and valid predictors for an array of social and economic outcomes. These subjective criteria are best assessed by means of subjective measures and thus in form of subjective indicators.

Typical subjective indicators in social and economic survey

In social and economic surveys there are a number of subjective indicators which are assessed rather regularly. Instead of providing a complete list, we will give a short overview of the most frequently assessed subjective indicators.

The most prominent criterion for a subjective indicator is *satisfaction* or *subjective well-being*. This can be either operationalized as satisfaction with life in general or as satisfaction with specific aspects of one's life, such as work, marriage or living conditions. Satisfaction is regularly assessed in numerous studies, such as the German Socio-Economic Panel (GSOEP), the European Social Survey (ESS), the German Welfare Survey, the European Value Study and the World Value Study.

The counterpart to satisfaction is *worries*. Worries are often assessed in the field of job security or health. Thus, worries are usually assessed based on specific aspects of the respondent's life. Studies that regularly assess worries are again the GSOEP and the German Welfare Study.

Another often measured complex of subjective indicators in social surveys are measures of *trust*, such as trust in the government, in the democracy etc. or general trust in others. Various questions regarding trust have been asked in the German General Household Survey (ALLBUS), in the Eurobarometer, in the International Social Survey (ISSP) and in the World Value Study (WVS).

The assessment of the most prominent subjective indicators

Usually satisfaction, worries and trust are measured by single item indicators. Typical phrasings of these questions are given in Table 1. Respondents are asked to rate their degree of satisfaction, worries, or trust as outlined in the given item formulation on a Likert-scale ranging from "not at all [satisfied, worried, trusting]" to "fully [satisfied, worried, trusting]".

Research, however, has shown that psychometric properties of multi-item scales, as used in psychology for example, are significantly higher than for single-item measures. In multi-item scales, item responses are aggregated to a sum or mean score of the underlying construct. In general, multi-item scales should be used the more the broader the construct of interest is, e.g. general life satisfaction vs. satisfaction with income. Broader constructs cannot be validly assessed by only one indicator.

Table 1. Typical item phrasings for measures of satisfaction, worries, and trust.

<p>General life satisfaction (WVS)</p>	<p>All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are “completely dissatisfied” and 10 means you are “completely satisfied” where would you put your satisfaction with your life as a whole?</p>
<p>Domain specific satisfaction (GSOEP)</p>	<p>How satisfied are you today with the following areas of your life? How satisfied are you with...</p> <ul style="list-style-type: none"> your health your sleep your job your housework your household income your schooling and professional education your place of dwelling your free time your family life the child care available?
<p>Domain specific worries (GSOEP)</p>	<p>What is your attitude towards the following areas - are you concerned about them?</p> <ul style="list-style-type: none"> General economic development Your own economic situation Your health Environmental protection Maintaining peace Global terrorism Crime in Germany Consequences of the expanding of the EU to the east Immigration to Germany Hostility towards foreigners or minorities in Germany <i>If you are employed:</i> Your job security?
<p>Trust in others (WVS)</p>	<p>Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?</p> <p>I 'd like to ask you how much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all?</p> <ul style="list-style-type: none"> - your family - your neighbourhood - people you know personally - people you meet for the first time - people of another religion - people of another nationality
<p>Trust in institutions (ALLBUS)</p>	<p>I am now going to read out a number of public institutions and organisations. Please tell me for each institution or organisation how much trust you place in it. Please use this scale.</p> <p>1 means you have absolutely no trust at all 7 means you have a great deal of trust</p> <p>You can differentiate your answers using the numbers in between. What about the –</p> <ul style="list-style-type: none"> Health service German constitutional court

	German Parliament Municipal administration Army Catholic church Protestant church Judicial system Television Newspapers Universities and other institutes of higher education German government Trade unions Police Job centres State pension system Employer associations?
--	---

Surveys, however, face tight time constraints. Therefore, the number of items used for multi-item scales needs to be restricted. In recent years, efforts have been made to construct ultra-short measures for constructs like personality or values to obtain multiple-item measures that can be used under the tight time restrictions of surveys. This approach could also be very fruitful for subjective indicators. On the one hand it would be a way to reduce lengthy item batteries, on the other hand broad constructs could be assessed more validly.

How can the subjective indicators be validated?

As mentioned above, subjective measures of subjective substances are hardly testable against their supposed ‘true scores’. The ratings represent feelings, evaluations or conditions inside the respondent that can hardly be perceived from an outside observer. Thus the respondent’s ‘true score’ remains a mystery.

There are, however, several means to investigate the quality of a self-report and to get an insight into the process of the judgment. If the individual’s response to a subjective indicator is reliable, respondents should answer the same question in the same way if asked again. To avoid memory effects when responding the second time a substantial time interval should be established between the two administrations, as it can be assumed that a judgment of one’s life satisfaction or of one’s job security is relatively stable over time, say across several weeks (given that no serious life event happens in that period of time). Therefore, one way of investigating the quality of a subjective indicator is to investigate its stability by administrating the same question to the same set of respondents twice. A measurement can be regarded as reliable if the retest stability, thus the correlation between the two administrations, is high.

Another way of investigating the reliability of a subjective indicator is to use multiple items for its assessment. As all items are assumed to measure the same underlying construct it can be assumed that all items are positively related to each other. The standardized intercorrelation among this set of items (Cronbach Alpha coefficient) could therefore be regarded as a measure of the measurement reliability.

Testing the validity of the measurement of a subjective indicator poses a larger challenge. Validity is usually investigated by comparing self-estimates to peers or expert ratings or by comparing it to an external (objective) criterion (e.g. self-rated health to results of biochemical tests or by the number of doctor's consultations). Such comparisons are difficult, if not impossible, if the substance rated is subjective and not objectively measurable and by others perceivable, like for example health. The only way of shedding some light on the validity of subjective indicators would be to investigate their construct validity. This could either be done – for multi-item indicators - by a factor analytic approach testing, if all items assumed to measure one indicator form, a common factor or in a multitrait-multimethod design testing several subjective indicators assessed by for example self-ratings and peer ratings.¹ Alternatively, this construct validation could be more theoretically driven by investigating hypothesized external correlations. For example it could be assumed that a person highly worried about his financial situation would be less likely to put high amounts of money in high-risk investments. This form of validation highly depends on plausible hypothesis. To develop such hypotheses we first of all need to learn more about the respondent's process of understanding the question and – most importantly – about the respondent's process of developing a judgment which is then given as an answer to the item.

One, at recent time very prominent, way of investigating the respondents' judgmental processes is the qualitative approach of cognitive interviews. In these cognitive interviews a small number of respondents (preferably rather heterogeneous with regard to their socio-demographic background) are asked to respond to the items in question. In doing that they are either invited to think aloud or to paraphrase specific terms or to explain why they chose a specific answer category. This approach can significantly increase our knowledge about the respondents' understanding, interpretation, and answering of a subjective indicator item. This becomes even more valuable, the less concrete an item is. Asking a respondent to think aloud while answering the item "Do you like to go to parties" – a typical Extraversion item – would probably not give us as much insight in the respondents mind as letting him think aloud about whether he is satisfied with life in general. This latter investigation could help inform us about

1 Peer ratings are also perceived as subjective indicators because they are a subjective judgment about a target person.

the cues like health, partnership, children, a secure job, or the financial situation, a person used to perceive him as either satisfied or dissatisfied. These cues can then be used to develop hypotheses for assumable correlations. For example, if all respondents refer first to their health as a cue for their level of satisfaction a significant positive relation between these two indicators can be assumed and tested in terms of validity. Health, however, will never be the only cue and cannot be regarded as a proxy for life satisfaction. Thus, a substantial but not maximal correlation can be anticipated.

Parallel to the development of cognitive interviews an alternative quantitative approach to shed light on the individual's judgmental process was developed (Jasso 2006). In this factorial survey method, respondents are asked to rate the level of a specified outcome variable (e.g. well-being or healthiness) based on a given fictitious description of a person and its characteristics (age, gender, income, eating habits, housing etc.). Respondents will be rating a large set of such fictitious descriptions, termed vignettes. Based on these ratings the implicitly used equations for assigning outcomes like well-being or healthiness can be retrieved using statistical techniques.

In sum, the quality of subjective indicators can be investigated by using methods such as test-retest reliability, Cronbach Alpha coefficients, construct validity in form of factorial structure or hypothesized correlations with external criteria and by use of cognitive interviews and vignette techniques.

Recommendations

Subjective indicators have been proven to possess predictive power for a large array of social and economic outcomes. For most of them only single item measures are used. The phrasing of the items and the response formats differ substantially across different surveys. The items used are in most cases not tested with regard to their psychometric properties. This however, can be done with very simple means. We therefore recommend the following:

- Subjective indicators should become more widely accepted and investigated as they are proven to possess a substantial predictive power.
- More ultra-short scale measures should be developed and validated
- In order to reach more comparability across studies and thus more comparable results and in order to profit most from validated measures, surveys should try to use the same form of measurement, i.e. the same item phrasings and the same response scales.

- Even though the validation of subjective indicators cannot as easily be conducted as for objective ones, we strongly recommend to investigate reliability and validity of these indicators in as much depth as possible.
- Further, we suggest investigating the respondent's judgmental process for the subjective indicators' measures by use of cognitive interviews and/or by the vignette technique. This allows to obtain a clearer picture of how the item is understood by the respondent and on which cues he bases his judgment.
- If results of cognitive interviews indicate that individuals strongly vary in their understanding of and in their way of responding to these single-item indicators researchers should examine whether it would be more fruitful and thus more valid to use multiple indicators which could be in turn less abstract instead.

References:

- Jasso, G. (2006): Factorial survey methods for studying beliefs and judgments. *Sociological Methods and Research*, (34), 334-423.
- Noll, H.-H. (2001): Subjektive Indikatoren: Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. In: *Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Ed.): Wege zu einer besseren informationellen Infrastruktur (ebk17)*. Baden-Baden.
- Veenhoven, R. (2007): Subjective measures of well-being. In: *McGillivray, M. (Ed.): Human Well-Being – Concept and Measurement*. Houndsmills, 214-239.