

Measures of Happiness: Which to choose?

Ruut Veenhoven¹

Chapter ?? in: Gaël Brulé & Filomena Maggino (Eds.) 'Metrics of Well-being', Springer, Dordrecht, 2016, pp. ??-?? DOI: ??

1 Introduction

Call for greater happiness

All humans want a satisfying life for themselves and their children and this appears in the high ranking given for happiness in the value hierarchy of students all over the world (Diener & Oishi 2004). Individually people seek ways to a more satisfying life and in Western societies this quest is manifest in the soaring sales of 'how-to-be -happy books', such as 'The art of happiness' (Dalai Lama & Cutler 1998). It is also reflected in the development of life-coaching businesses. Citizens in western societies also call on their governments for greater happiness and 85% of the British agree with the statement that 'a governments prime aim should be achieving the greatest happiness of the people, not the greatest wealth' (BBC 2006, question 14). Consequently, interest in happiness is rising among policy makers; happiness is a new topic on the political agenda, next to sustainability. A recent manifestation of this trend is the international conference on Happiness and Wellbeing held at the UN headquarters in New York in April 2012 (Bhutan 2012), which was followed in June 2014 by a decision of the general assembly to celebrate an 'International Day of Happiness' on March 20th every year.

Demand for facts

Calls for greater happiness are often accompanied by recommendations about ways to achieve that. At the individual level such advice typically involves 'alternative' ways of life, such as consuming less and meditating more, while at the political level greater happiness for a greater number is seen in social reform, such as less economic competition and more family life (e.g. Layard 2005). Yet a rational pursuit of happiness should be based on established fact rather than on ideological belief. As such the pursuit of greater happiness is similar to the pursuit of better health. In the past we have learned a lot from empirical research on conditions for good health and, using that information, we live now longer than ever before in human history. Likewise, orienting on scientific knowledge about happiness will probably mean that we will also live happier long lives.

¹ Erasmus Happiness Economics Research Organization, Erasmus University Rotterdam, Netherlands, www.eur.nl/ehero

Need for good measures

To be able to do so, some premises need to be met: (i) We need to know what happiness actually is, thus, we need a clear definition of this construct. (ii) We need measures that capture the defined concept of happiness validly and reliably. In this chapter I review the available measures of happiness in the sense of life-satisfaction, drawing on measures and findings gathered in the World Database of Happiness (Veenhoven 2016a). We will see that there are many acceptable measures of happiness. Hence a next question is which measures are most appropriate in what context and this question is also addressed in this chapter.

2 Concept of Happiness

The word 'happiness' is used to denote different *concepts*. In the widest sense it is an umbrella term for all that is good. In this meaning it is often used interchangeably with terms like 'wellbeing' or 'quality of life' and to denote both individual and social welfare. Here the word happiness is used in the more limited sense of subjective satisfaction with life. Limitation to a clear meaning is required for a fruitful review of measures of happiness.

2.1 Definition of happiness

In this chapter I follow the definition of happiness which underlies the World Database of Happiness. Overall happiness is the degree to which an individual judges the overall quality of his/her own life-as-a-whole favorably. In other words: how much one likes the life one lives. The key terms in this definition are explained in Veenhoven (1984, 2016b).

2.2 Components of happiness

When evaluating the favorableness of our life, we tend to use two more or less distinct sources of information: our affects and our thoughts. One can decide that one feels fine most of the time and one can also judge that life seems to meet ones, conscious demands.

These appraisals do not necessarily coincide. We may feel fine generally, but nevertheless be aware that we failed to realize our aspirations, or reversely, we may have surpassed our aspirations, but nevertheless feel miserable. Using the word 'happiness' in both these cases would result in three different kinds of happiness, the overall judgment as described above and these two specific appraisals. I refer to the encompassing judgment, the core concept, as *overall happiness*. A synonym for overall happiness is 'life satisfaction'. The components are referred to as *hedonic level of affect*, the affective component, and *contentment*, the cognitive component (Veenhoven 2009). The distinction between 'affective' and 'cognitive' aspects of happiness is commonly made in the

literature, but seeing them as 'components' of overall happiness is not (Kainulainen & Veenhoven 2016).

Hedonic level of affect

Hedonic level of affect is the degree to which various affects that someone experiences are pleasant in character. The concept of hedonic level concerns only the pleasantness experienced in affects, that is, the pleasantness in feelings, in emotions, as well as in moods. So a high hedonic level may be based on strong but passing emotions of love, as well as on moods of steady calmness. A person's average hedonic level of affect can be assessed over different periods of time: an hour, a week, a year, as well as over a lifetime. The focus here is on 'characteristic' hedonic level. That is so to say: the average over a long time-span such as a month or a year. The concept does not presume subjective awareness of that average level.

Contentment

Contentment is the degree to which an individual perceives his/her aspirations are being met. The concept presupposes that the individual has developed some conscious wants and has formed an idea about their realization. The factual correctness of this idea is not at stake. The concept concerns the individual's subjective perception.

2.3 Difference with related concepts

This concept of happiness can be further clarified by noting the differences from related notions. Below I will first distinguish satisfaction with life from other qualities of life and then distinguish life-satisfaction (happiness) from other satisfactions. Note: many these different concepts are often called by the same name.

2.3.1 Difference with other qualities of life

The term 'quality of life' suggests that all merits can be integrated in one final scale of worth. This is not the case. The term is merely an umbrella for different notions of what is good with respect to one's life. Below I delineate four qualities of life and show that happiness fits only one of these.

Quality-of-life concepts can be sorted using two distinctions, which together provide a fourfold matrix. The first distinction is that between *chances* and *outcomes*, that is, the difference between opportunities for a good life and the good life itself. A second difference is found between *outer* and *inner* qualities of life, in other words between external/ and internal features. The combination of these two dichotomies yields a fourfold matrix. This classification is presented in scheme 1.

Scheme 1: **Four qualities of life**

	<i>Outer qualities</i>	<i>Inner qualities</i>
<i>Life-chances</i>	Livability of environment	Life-ability of the person
<i>Life-results</i>	Usefulness of life	Satisfaction

Source: Veenhoven 2000

Livability of the environment: the left top quadrant denotes the meaning of good living conditions, in short ‘livability’. Livability is not what is called happiness here. It is rather a precondition for happiness and not all environmental conditions are equally conducive to happiness.

Life-ability of the person: the right top quadrant denotes inner life-chances. That is: how well we are equipped to cope with the problems of life. Sen (1992) calls this quality of life variant ‘capability’. Ability to deal with the problems of life will mostly contribute to happiness as defined here, but having capability is not identical to being happy.

Usefulness of life: the left bottom quadrant represents the notion that a good life must be good for something more than itself. This is often referred to as ‘meaning of life’. A meaningful life is not necessarily a happy life, one can sacrifice one’s happiness for a good cause.

Subjective satisfaction: finally, the bottom right quadrant represents the inner outcomes of life. That is the quality of a life in the eye of its beholder. As we deal with conscious humans, this quality boils down to subjective satisfaction with life. This is commonly referred to by terms such as ‘subjective wellbeing’ and the concept of happiness as defined above fits this category. Still not all satisfaction is ‘life-satisfaction’, as we will see below.

2.3.2 *Difference with other satisfactions*

Even when we focus on subjective satisfaction with life, there are still different meanings associated with the word happiness. These meanings can also be charted in a fourfold matrix. In this case, the classification is based on the following dichotomies: Life-aspects versus life-as-a-whole and passing delight versus enduring satisfaction.

Scheme 2: Four kinds of satisfaction

	<i>Passing</i>	<i>Enduring</i>
<i>Part of life</i>	Pleasure	Domain-satisfaction
<i>Life-as-a-whole</i>	Peak-experience	Life satisfaction

Source: Veenhoven 2015

Pleasure: the top-left quadrant represents passing enjoyments of life-aspects. Examples would be delight in a cup of tea at breakfast, the satisfaction of a chore done or the enjoyment of a piece of art. The concept of happiness used here is broader however. It concerns both overall satisfaction and life-as-a-whole. Though fleeting enjoyment obviously contributes to a positive appreciation of life it is not the whole of it.

Domain satisfaction: the top right quadrant denotes enduring appreciation of life-aspects, such as marriage-satisfaction and job-satisfaction. Domain-satisfactions are often denoted with the term happiness: a happy marriage, happy with one's job, etc. Yet here the term happiness is used in a broader sense, not for the satisfaction with aspects of life, but for the satisfaction with life-as-a-whole.

Top-experience: the bottom left quadrant denotes the combination of passing experience and appraisal of life-as-a-whole. That combination occurs typically in top-experiences, which involve short-lived but quite intense feelings and the perception of wholeness. Again this is not the kind of happiness aimed at here. A moment of bliss is not enduring appreciation of life.

Satisfaction with one's life-as-a-whole: the bottom-right quadrant represents the combination of *enduring* satisfaction with one's *life-as-a-whole*. This is how I define happiness. A synonym is 'life-satisfaction'.

2.4 Focus on 'present' happiness

Evaluations of one's life as a whole may concern different periods of life: earlier life, current life and (expected) future life. In this chapter I restrict to evaluations of 'present' life.

2.5 Currency of this concept

Many things are called 'happiness'. The meaning addressed here is quite often used in present day social science research and the definition dominates in happiness economics and sociological studies. In positive psychology it is sometimes called 'hedonic happiness' and distinguished from 'eudaimonic happiness'². As we will see in section 3.2 about half of the measures denoted 'happiness' fit *this* concept of happiness as life-satisfaction.

3 Measures of Happiness

3.1 Techniques

Happiness as defined above is commonly measured using self reports. In the case of hedonic level, rating by others is also possible.

Self report

Happiness as defined above is something on our mind and can for that reason be measured using self-reports. In this respect happiness differs from many concepts in psychology that do not require subjective awareness, e.g. 'neuroticism', neurotics mostly do not know how difficult they are, since excessive ego-defense is part of the syndrome. Hence neuroticism is mostly measured using observations of symptoms from which an expert infers the degree of neuroticism. These symptoms are often measured using self reports and as such self-reports can be used to measure neuroticism indirectly. Direct questions such as 'How neurotic are you?' are not appropriate for this subject.

Direct self-report is possible in the case of happiness and is the only suitable technique in the cases of 'overall happiness' and 'contentment'. Happiness cannot be measured using peer ratings; friends cannot look into your head and can at best guess how much you like your life.

Self reports are typically made in response to single direct questions, which respondents answer by ticking one of several pre-given answer options. Examples of such questions are presented in section 3.3. Often several such questions are combined in a questionnaire, such as the often used five item Satisfaction with Life Scale (SWLS) of Pavot & Diener (1993). Self reports can also be derived from content analysis of open interviews or ego-documents.

Rating by others

The case of hedonic level is different. As noted above, this concept does not require awareness of average affect, one can feel pleasant most of the time without being

² 'Eudaimonic' happiness is another word for 'positive mental health' and covers various traits deemed beneficial, such as autonomy, identity and meaningfulness. Most of the meanings addressed by this term fit the right top quadrant of scheme 1 It is a fuzzy concept and several of its constituent cannot be adequately measured using self-reports.

conscious of that, as is typically the case with children. Hence, hedonic level can be measured using time sampling techniques in which an investigator aggregates repeated self reports of momentary affect. Since hedonic level reflects in non-verbal behavior it can also be measured using observation techniques, such as time-sampling of smiling, and using ratings by parents or peers.

No physiological indicators

Dependable biological indicators of happiness are not available as yet and probably never will be. Though one's experience of happiness draws on a biological substrate it will be difficult to infer the signal from the machinery.

No qualitative measures

All these measures quantify how *much* people like the life they live. This is typically not assessed in qualitative studies on happiness (listed in section [A109](#) of the Bibliography of Happiness), which focus on *what* people like or not. A few studies have estimated the degree (quantity) of happiness on the basis of content analysis of qualitative data such as life review interviews and ego-documents (Newman & Langner 1981, Danner et al 2001).

3.2 Validity

'Validity' in science is the correspondence between concept and measurement. In psychology this is often difficult to establish, since fuzzy concepts prevail in this discipline. Rather than assessing correspondence with a clear concept, psychologists typically infer conceptual meaning from inter-correlations between items in their questionnaires (construct validity) and correspondence with related constructs (convergent validity). In the case of happiness we can do better. Since we have a well-defined concept, we can check directly whether the questions used to elicit a self-report reflect the concept of happiness as defined above. This is called testing for 'face-validity'.

Selection on face-validity

I have screened all the measures of happiness ever proposed in the scientific literature. I did this in the context of my World Database of Happiness, which is an online archive of research findings on happiness. Since this collection limits to happiness as defined above, I had to weed out findings based on measures that tap slightly different phenomena.

To date I have inspected some 5000 measures. About half of these appeared to tap essentially different things, such as the Oxford Happiness Inventory (Hills & Argyle 2002), many items of which are about mental health and tap aspects of life-ability, depicted in the upper right quadrant of scheme 1. The much used Satisfaction With Life Scale: SWLS (Pavot & Diener 1993) did not pass this test either because one of its five

items is also about something other than happiness³.

Currently⁴, 2,192 measures have passed this check for face-validity, most of these are single questions (1.557) that differ slightly in wording. Next there are multiple question inventories (458), and observation schedules for assessing behavioral manifestations of hedonic level of affect (99).

Collection of accepted measures

Acceptable measures are included in the collection ‘Measures of Happiness’ of the World Database of Happiness (Veenhoven 2016c). In this chapter I draw on this collection and on the findings yielded with accepted measures of happiness stored in the findings archive. I use the World Database of happiness as an online electronic supplement to this chapter. Note: I insert links to the database in this text, which the reader can use to see more detail than this book chapter can provide. An additional advantage of this technique is that it provides the reader with information added to the database after this text was printed.

3.3 Differences in valid measures of happiness

The main aim of the World Database of Happiness is to prepare research findings on happiness for synthetic analysis and this often requires comparison of findings obtained with the same measure of happiness. For this reason all accepted measures in the WDH are classified by both their substantive meaning and the method of assessment used. Since subtle differences still remain a further classification by sub-variant makes it possible to identify fully identical measures of happiness. The classification of happiness measures is shown on scheme 3. The links provide access to detail; use Click+Control. An example of each category is presented below.

Scheme 3: Classification of happiness measures

Substantive Meaning	Focus	The kind of happiness addressed.
	Time-frame	The period considered
Method of assessment	Mode	The technique by which happiness is assessed
	Scale-type	How the observation is scored
	Scale-range	Number of degrees of happiness distinguished
Sub-variant	Wording	Variation in phrasing of otherwise equivalent questions

³ The last item in the SWLS is a question whether one would live the same life if one could live one’s life over again. The assumption is that happy people will want to repeat their life. Yet one can enjoy present life but still be open for something else.

⁴ As assessed on August 30, 2016.

The Kind of Happiness Addressed

Three kinds of happiness are distinguished above: 1) Overall Happiness, also called 'life-satisfaction, 2) an affective component called 'Hedonic level of affect' and a cognitive component called 'contentment'. All these kinds of happiness can be measured using single direct questions, such as the examples below. The links to the right lead to full detail in the World database of Happiness.

- Overall Happiness:
["How do you feel about your life as a whole...?"](#)
- Affective Happiness Component:
["How is your mood these days...?"](#)
- Cognitive Happiness Component:
["How do you feel about what you are accomplishing in life...?"](#)

Some question cover more variants of happiness simultaneously and these are classified as 'mixed measures'. An example is:

- Mixed Measures:
["How many days in the previous week did you feel happy?"](#)

Time Frame

Measures of happiness differ in the period addressed, the most commonly used time reference is these days'. Some examples are:

- Momentary, Now:
["How are you feeling now...?"](#)
- Yesterday:
["Overall, how happy did you feel yesterday?"](#)
- Last Year:
["Generally, how happy have you been with your personal life during the past 12 months?"](#)
- Currently:
["How satisfied are you with your life as a whole these days? Would you say you are...?"](#)

Observation technique

As noted in section 3.1 different techniques are used for measuring happiness: Some examples are:

- Self report:
[In thinking over the past year, indicate how elated or depressed, happy or unhappy you have felt in the last year?"](#)
- Rating by Others:
["Overall how does your child usually feel?"](#)

Response scale

Answers to questions can be rated on different scales. Examples:

- Verbal Scales:
terrible, unhappy, mostly dissatisfied, mixed, mostly satisfied, pleased, delighted
- Numerical Scales:
1 2 3 4 5 6 7 8 9 10
dissatisfied satisfied
- Graphical Scales: "7-points-smiley-scale: from sad face to happy face here

In the following table 1 I tabulate how many variants of each of these types of measures have been used. For example, the term 'happiness' has been used in 406 measures, mostly single direct questions that differ in timeframe, response scale and wording.

Table 1
Variations in acceptable measures of happiness

<i>Characteristics</i>	<i>Variants of this kind</i>
Kind of happiness addressed	
Overall Happiness	
- Keyword 'happiness'	406
- Keyword 'life satisfaction'	641
- Keryword 'life quality'	21
Affective component: Hedonic level	
- Average affect: overall estimate	72
- Average affect; sumscore	64
- Affect balance	244
- Cheerful appearance	37
- Cheerful person	29

Cognitive component: Contentment	
- Realization of wants	65
- Best-Worst possible life (Cantril ladder)	43
Mixed Measures	329
Time Frame	
Momentary, Now	158
- Last Instant	52
- Last Hour	8
- Last Part of the Day	7
- Last Day	44
Yesterday	57
Currently (Presently, Today, these Days)	827
- Last Week	106
- Last Month, Last Few Weeks	117
Last Quarter	14
Last Year	48
Last Years	5
Over Lifetime	8
Currently (Presently, Today, these Days)	532
Generally	318
Hitherto	48
Since Event	9
Various Time Frames (in case of mixed measures)	77
Time Frame Unspecified	509
Observation Technique	
Self-Report	
- Single question, once	1557
- Single question, repeated	33
- Multiple questions; once	534
- Multiple questions repeated	56
Rating by Others	
- Intimates (family, peers)	37
- Teachers, nurses	18
- Interviewer	10
- Clinician	21
Focussed interview	26
Content-analysis eego-documents	3
Time sampling	13
Response scale	
Verbal Scales	1,288
Numerical Scales	556

Graphical Scales

-	
-	
- Faces scale	33
- Ladder scale	58
- Life graph	8
- Open line scale (VAS)	46

* As assessed on August, 2016.

4 Strengths and weaknesses

All the measures considered here have passed a test for face-validity, so all are 'valid' in the sense that they concern the concept of happiness as defined in section 1 and not something else. This is not to say that they all the measures measure the concept equally well. The measures may differ in vulnerability to distortions and in precision. Below I summarize the main findings on the psychometric properties of the various measures. I draw on the literature gathered in the Bibliography of Happiness (Veenhoven 2016d), section Ca Measurability of Happiness, which includes some 400 publications.

This literature has several limitations. A main limitation is that many publications deal with the measurability of happiness in general, while the question is rather which variant is the most applicable in a particular situation. For example: a single question that works well in the general public may not be understood by demented residents in a care home. A second limitation is that much of this literature is aimed at demonstrating that a particular kind of measurement error exists, such as social desirability bias, while remaining silent about the size of the distortion and the degree to which different sources of error add up or balance out.

In the review below I limit to self-report measures of happiness. Discussion of observation schedules of behavioral manifestations of hedonic level would be speculative, since there is little psychometric research in this matter. I mention some main issues and provide links to the literature, again using links to relevant sections in the Bibliography of Happiness.

Comprehensibility for respondents

Do respondents understand the questions? An indication of understandability is how many of them tick the 'Don't know' option when presented with the response scale or who simply skip the question (No Answer). These responses appear to be quite rare, in surveys of the general population typically less than 1% of the respondents is found in these categories. Literature on this issue is listed in section Ca01.04 of the Bibliography of Happiness ['Do people respond to questions about happiness?'](#)

Vulnerability for distortions

Several studies have revealed specific distortions in responses to questions about happiness, such as these due to context of the interview, characteristics of the interviewer and day of the week. Literature on that matter is gathered in the Bibliography of Happiness, Ca02 [Reliability of happiness measurements](#).

Some of these distortions are random and do not influence the central tendencies in big samples and nor do they affect the direction of correlation, though they will reduce the size. This problem can be solved using a 'disattenuation' technique (Headey & Wearing 1992). Systematic bias is more problematic, e.g. higher ratings in face-to-face interviews than in web surveys. Such distortions can be corrected for when identified. Literature on correction methods is listed in the section on [Bias correction](#) of the Bibliography of Happiness.

Correspondence with ratings by others

Self ratings of happiness correspond typically satisfactory with ratings made by people who know you well, such as by friends, parents and teachers. Findings on this issue are gathered in section H9 [Reputation of Happiness](#), in the collection Correlational Findings of the World Database of Happiness. Note: the correspondence is not perfect, for example parents tend to overrate the happiness of their children.

Stability over time

Over-time stability of happiness is has been assessed in some 100 follow-up studies, the results of which are gathered in section H5.2 [Actual happiness career](#) of the WDH collection correlational findings. Correlation with baseline after one year is typically around +.60 and drops gradually to +.30 in 20 years. There are different reasons for this low repeat correlation. One reason is the imprecision of the happiness measures used; a respondent who remains equally happy may tick '7' one year and '6' the year after and the above mentioned distortions in the response process may work out differently at different observations. Another reason is a real change in happiness, due to over time variation in living conditions and life ability. The longest large-scale follow-up study available at the moment was recently published by Headey and Muffels (2016).

Differentiation across populations

In spite of these limitations, happiness measures show clear differences between populations, both across nations and within nations. The rich data on cross-national differences are listed in the WDH section [Happiness in Nations](#). An illustrative finding is that in 2007 average happiness differed almost seven points on scale 0-10 between Togo (1,6) and Costa Rica (8,5). Differences in happiness within nations are typically smaller, but still sizable, such as those between single and married persons and people in good and bad mental health. See WDH section [Happiness in Publics](#).

Sensitivity to conditions for happiness

A common issue in social indicators research is the low correlation between subjective and objective indicators of the same, such as in the case of crime, where there is often little correspondence between fear of crime and actual crime rates. In the same vein a point of discussion in happiness research is the low correlation of happiness with some factors deemed conducive to it, such as [education](#). Such unexpected, and often unwelcomed, results are sometimes attributed to the poor quality of the happiness measures used. This is possibly correct in particular cases, but not correct as a rule. Happiness correlates strongly with several other factors, such as with unemployment and predicts how long one will live. See respectively the WDH sections E2 on happiness and [employment](#) and section P6.1.4 on [later longevity](#).

Comparability across nations

There are doubts about the reality value of the differences in average happiness mentioned above and one of qualms is that they are largely due to cultural measurement bias, such as difference in understanding of questions and response tendencies. One such source of bias is cultural variation in aptness to tick the highest option on 0-10 numerical response scales (Brule & Veenhoven 2016). The rich literature on this topic is gathered in section Ca03.02 [Comparability across nations/languages](#) of the WDH Bibliography of Happiness.

There is no doubt that cultural measurement bias exists, the question rather is how big this bias is. Veenhoven (2012, section 4.3) estimates that about 5% of the variation in average happiness across countries is due to cultural measurement bias

5 How to choose a measure of happiness

As we have seen in section 3, more than a thousand valid measures of happiness exist. How should a researcher choose among this multitude? The first step is to select the happiness variant that fits the research question best, the second step is to select on methodological qualities and the last step is to choose on the bias of pragmatic considerations, such as cost.

5.1 Conceptual considerations

Happiness variant

Though all measures of happiness considered in this chapter concern the subjective enjoyment of one's life as a whole, there are still differences in focus within this conceptual realm. In section 2.2 we distinguished between *overall happiness*, i.e. life satisfaction, and its two 'components', the affective component called *hedonic level* and the cognitive component called *contentment*. The collection of Happiness Measures is classified accordingly, all measure codes start with either A for Affect, C for Contentment, O for Overall happiness or M for Mixed measures.

Which variant to choose depends on the research question in the first place; a measure of contentment may be most appropriate if the question is how exposure to advertising works out on subjective wellbeing. The choice of variant depends also on the population investigated, in the case of young children hedonic level is the only option, since youngsters have not yet developed a view on their life as a whole, which is required to answer questions about overall happiness and contentment.

The choice of a happiness variant has also methodic consequences; the use of other ratings is possible only in the case of hedonic level.

Time frame

Another conceptual consideration is the time frame of the measure, the choice of which will also depend on the research question. If the question is about the livability of a society, one would like to know how happy its inhabitants generally feel. If the question is how people adapt to divorce, it may be more apt to focus on how happy they felt since that event. The choice will again also depend on the chosen population. Demented people can report how happy they feel in the moment, but typically not how happy they have felt in the last year.

As we have seen above in table 1, some 500 measures do not specify a time frame. I cannot imagine situations in which such unclarity is useful and for that reason advice against the use of such measures.

5.2 Methodological considerations

The many available measures of happiness differs also in the observation techniques and rating scales that are used. What is the best choice of measure for what situation?

Single or multiple questions?

Most entries in the WDH collection of Happiness Measures are single direct questions (71%). See above scheme 3. One of the reasons for this is that several multiple item questionnaires failed my test for face validity, because one or more of the questions did not fit the concept of happiness as defined above in section 2. Several of these questionnaires ask how happy are compared to others, rather than how happy you feel yourself. Such questions are rejected. Even if one is happier than one's neighbour is, one can still be unhappy. This invalidates the four-item 'Subjective Happiness Scale' (SHS) of Lyubomirsky and Lepper (1999), the second item of which is perceived happiness relative to peers. For the same reason the question 'I have been happier than I am now' is deemed unacceptable. Being less happy than before does not imply that one is unhappy. This is not just a matter of unfortunate formulation, but rather a matter of conceptualization. The designers of such questionnaires typically think of happiness as a fuzzy concept and measure it using proxies instead of addressing it by itself. This appears in their use of words: in their theoretical accounts they speak of 'subjective wellbeing' or

'psychological well-being' and use the terms 'happiness' and 'life-satisfaction' only in the name of the scale. Still there are several multiple item measures that did pass my test, for instance the 4-item variant of Diener's Satisfaction With Life Scale.

An advantage of single questions is that it clear what is measured; they have high construct validity; at least when the question is clearly formulated. An evident practical advantage is that they require less interview time.

An advantage of using multiple questions is that it can reduce measurement error, among other things because subtle differences in the interpretation of words will balance out. If spread out in an interview the use of multiple questions can also reduce contextual bias. Andrews & Withey (1976) did that by asking the same single question at the beginning of an interview (Life1) and at the end (Life2), and then using the mean (Life3) of the responses.

What response scale?

As can be seen in table 1, there is much variety in response scales used. What is the best choice from a psychometric point of view? Several issues need consideration. Literature on this matter is found in subject section Cb01.01.02 [Rating scales](#) of the Bibliography of Happiness among which a recent review by Ludwigs & Henning (2016).

Even or uneven: The first issue consists of deciding whether to use an uneven number of scales levels, containing a midpoint, or an even number of scales levels, forcing respondents to take a position. Several researchers believe the first option is to be preferred against the second, the main argument is that omitting a midpoint means that participants randomly choose one of the scale points that are closest to the midpoint (e.g. Krosnick & Presser, 2010).

Number of options: The second challenge is to choose the right number of answer options. In theory the probability of a match between one's happiness and a given answer option is greater the more options available, but too many options can cause ambiguity which then can lead erratic responding. The research literature suggests that 5 or 7-point rating scales are the best, yet below, in the discussion of pragmatic considerations, we will see that 0-10 numerical scales are more commonly used and thus provide more opportunities for comparison.

Labeling of response options: Another issue concerns full labeling vs. a partial labeling of response options. In the case of full labeling each option is described using words, such as 'very happy', 'pretty happy' and 'not too happy'. In the WDH these are classified as 'verbal scales'. In table 3 we have seen that this is the most common technique. In the case of 'partially labeled' response scale, the response options are denoted with numbers for which only the highest and lowest are given descriptive words, e.g. '10 completely happy 9 8 7 6 5 4 3 2 1 0 complete unhappy'. Sometimes also a middle option is also defined with a word, such as '5 neutral'. In the WDH this variant is called a 'numerical

scale’.

Verbal response scales are the most commonly used scales. Arguments for the use of verbal scales are that respondents prefer them and that such scales produce higher reliability and validity scores because the labeling of each scale point facilitates understanding for respondents.

Yet there are also disadvantages to this technique. One is that it limits the number of response options, verbal response scales typically present 3 to 5 options; 10 options is typically too much to denote using words. Another disadvantage is that respondents may differ more in their interpretations of words, than of numbers, especially in countries where numbers are used for grading of school performance; where e.g. the number 7 on a 0-10 scale would be more clear than the verbal label ‘pretty happy’. For the same reason numerical scales allow better comparability across nations; the word ‘heureux’ in French may denote a higher degree of satisfaction than the word ‘happy’ in English, but the number 7 is likely to have the same valence in both nations. Lastly numerical scales allow more possibilities for statistical analysis, because they come closer to the demand for equidistance between response options.

Another point of discussion is whether to add numbers to the verbal labels of full labeled response scales. Experiments on this topic show that people answer differently to scales containing the same verbal labels but diverging numerical labels (0 to 10 vs. -5 to +5) indicating that respondents use numerical labels “to disambiguate the meaning of scale labels, resulting in different interpretations and, accordingly, different subjective scale anchors” (Schwarz et al., 1991, p. 570).

Agree/disagree format: Some investigations present their respondents with a statement, such as ‘I am a happy person’ and ask respondents to what extent they agree or disagree. This response format appears to be vulnerable for acquiescence distortion, such as the tendency for the respondent to respond politely rather truthfully.

Graphic scales: A last issue concerns the use of graphic response scales for rating degree of happiness, using pictures rather than words or numbers.

One such technique uses an open line on which only the extremes are denoted using words, such as ‘happy’ and ‘unhappy’. Respondents indicate their happiness by marking a point on that scale, on electronic devices typically using a slider. Such ratings are later transformed in a number. This is called ‘open line scale’ or Visual Analog Scale (VAS). In theory such scales allow more precise ratings, in practice they show little gain in precision, the absence of defined reference points on the scale invites respondents to make rough responses and when the data is crunched, results are typically reduced to a 10 step scale by the researcher.

A wordless, numberless technique for rating degree of happiness is to use smilies to indicate response options. The smilies use facial expressions that are universally recognized as expressing degrees of happiness. This marks a main advantage of this method; it allows cross-cultural comparisons. The method is also suited for respondents

who are illiterate. A limitation of these rating scales is that they focus the respondent's attention on emotional experience and as such they are better suited for measuring hedonic level than overall happiness or contentment.

5.3 Pragmatic considerations

Research often involves comparison with results of earlier studies, e.g. when the aim is to assess progress over time or differences from other samples of the same population. Comparison requires the use of comparable measures and ideally identical ones. Hence researchers would be wise to choose a measure of happiness that has been used earlier in studies with which they want to compare their results. The WDH collection of Happiness Measures is useful for this purpose, since it allows selection of measures by populations in which these have been used.

An issue related to comparability of rating scales in particular is the possibility to transform observations made using one particular scale to another scale, typically from different verbal response scales to a common 0-10 numerical scale. Such transformations can be made in several ways and new techniques have been developed recently (DeJonge 2016). One of these techniques is the International Scale Interval Method (Kalmijn et. al 2011), the application of which requires that the valence of the response options in a question in a particular language are rated by native judges. Another technique is the 'Reference Distribution Method' (DeJonge et. al 2014), which requires that a distribution of responses to another question on the same kind of happiness in the same time and population is available. Whether such data is available can be checked in the WDH.

Investigators should aim to use measures with good psychometric properties and the available information on that matter is included in the WDH collection of Happiness Measures: reliability indicators are data on over time correlations and, in the case of multiple item measures, also inter-item correlations. Differentiating power of measures of happiness appears in the correlations yielded using them.

As indicated above, the best comparable measure is not always psychometrically the best one. In such cases investigators have to settle for the less than ideal. Using the collection of Happiness Measures allows investigators to do so on the basis of full information.

6 Summary

Of the many measures claimed to assess 'happiness' about half tap something else than happiness as 'the subjective enjoyment of one's life as a whole'. The measures that do fit this concept are gathered in the 'Collection of Happiness Measures' of the World Database of Happiness. To date this collection contains about 2000 measures, most of

which are single direct questions that differ slightly in phrasing and response format.

The strength of each of these selected measures is that it is clear what they measure; there is no doubt about their validity. A strong point of the WDH collection is that investigators can choose a variant that fits their research needs best.

One weakness of most of these measures is that they are not very precise. While corrections can reduce this reliability deficit, they cannot entirely solve it. A further weakness of this family of indicators is that the many small differences between measures reduce comparability across studies. New transformation techniques will improve comparability, though it will remain a problem in happiness research.

REFERENCES

Andrews, F.M. & Withey, S.B. (1976). *Social Indicators of Well-being: Americans' Perceptions of Life Quality*. New York, USA: Plenum Press.

BBC (2006), *The happiness formula*. GfK-NOP poll 421059. Available at: http://news.bbc.co.uk/2/hi/programmes/happiness_formula/4771908.stm

Bhutan (2012) *Defining a new economic paradigm. Report of the high-level meeting on well-being and happiness*, 2 April 2012, United Nations headquarters, New York.

Available at:

<https://sustainabledevelopment.un.org/index.php?page=view&type=400&nr=617&menu=35>

Brule & Veenhoven (2016) The '10-excess' phenomenon in responses to survey questions about happiness, *Social Indicators Research*, online since February 2016

DeJonge, T.; Arends, L.R.; Veenhoven, R (2014) Homogenizing responses to different survey questions on the same topic. *Social Indicators Research*, 117, 275 - 300

DeJonge, T., Kalmijn, W. & Veenhoven, R. (2016) *Dealing with diversity in survey questions on the same subject, Improving comparability across surveys and nations*, Springer, in press

Dalai Lama & Cutler, H (1998), *The art of happiness, A handbook for living* Riverhead Press, USA.

Danner, D.D., Friessen, W.V. & Snowdon, D.A. (2001) Positive emotions in early life and longevity: Findings from the Nun Study. *Journal of Personality and Social Psychology*, 80, 804 - 813

Diener, E. & Oishi, S. (2004) *Are Scandinavians happier than Asians? Issues in comparing nations on subjective wellbeing*. In: Columbus, F.; Ed.: "Asian Economic and Political Issues", 20: 1 - 25

Headey, B. & Wearing, A.J. (1992) *Understanding happiness: A theory of subjective well-being*, Longman Cheshire, Melbourne, Australia

Headey, B. & Muffels, R. (2016) *Towards a theory of life satisfaction Accounting for stability, change and volatility in 25-year life trajectories in Germany*, IZA Discussion Paper No. 10058, Institute for the Study of Labour, Bonn Germany

Hills, P. & Argyle, M. (2002). The Oxford Happiness Questionnaire: A compact scale for the measurement of psychological well-being. *Personality and Individual Differences*, 33, 1073 -1082

Kalmijn, W.M. Arends, L.R. & Veenhoven, R. (2011) Happiness scale interval study: Methodological considerations, *Social Indicators Research*, 102: 497-515

Krosnick, J. A., & Presser, S. (2010). *Question and questionnaire design*. Handbook of Survey Research, 2, 263-314.

Kainulainen & Veenhoven (2016) *Life satisfaction is more a matter of feeling well, than having what you want*, EHERO working paper, Erasmus University Rotterdam

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

Ludwigs, K. & Henning, L. (2016). *How to measure subjective well-being – A practical review*. HRO Working Paper, Happiness Research Organization, Dusseldorf, Germany.

Lyubomirsky, S. & Lepper (1999), A Measure of Subjective Happiness: Preliminary Reliability and Construct Validation, *Social Indicators Research*, 46, 137 – 155

Newman, H.M. & Langner, E.J. (1981) Post-Divorce Adaptation and the Attribution of Responsibility, *Sex Roles*, 7, 223 - 232

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment*, 5(2), 164-172.

Saris, W.E., Revilla, M., Krosnick, J.A., & Shaeffer, E.M. (2010). Comparing questions with Agree/Disagree response options to questions with item-specific response options. *Survey Research Methods*, 4, 61-79.

Schwarz, N., Knäuper, B., Hippler, H., Noelle-Neumann, E., & Clark, L. (1991). Rating scales: Numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55, 570-582.

Sen, A: (1992), *Capability and wellbeing*, In: Sen, A. & Nussbaum, M (Eds.) 'The quality of life' Clarendon Press, Oxford, UK, pp. 30-53

Veenhoven, R. (1984). *Conditions of Happiness*, Reidel (now Springer), Dordrecht, Netherlands

Veenhoven, R. (2009) *How do we assess how happy we are?* In: Dutt, A. K. & Radcliff, B. (eds.) 'Happiness, Economics and Politics: Towards a multi-disciplinary approach', Edward Elger Publishers, Cheltenham UK, ISBN 978 1 84844 093 7, Chapter 3, page 45-69

Veenhoven, R. (2012) Cross-national differences in happiness: Cultural measurement bias or effect of culture? *International Journal of Wellbeing*, 2: 333-353

Veenhoven, R. (2016a). *World Database of Happiness*, Erasmus University Rotterdam, the Netherlands. Available at: <http://worlddatabaseofhappiness.eur.nl>

Veenhoven, R. (2016b) *Bibliography of Happiness*, World database of Happiness, Erasmus University Rotterdam.

Available at: http://worlddatabaseofhappiness.eur.nl/hap_bib/bib_fp.php

Veenhoven, R. (2016c) *Concept of Happiness*, Chapter 1 in 'Introductory text to the collection of Happiness Measures', World Database of Happiness, Erasmus University Rotterdam.

Available at: http://worlddatabaseofhappiness.eur.nl/hap_quer/introtext_measures2.pdf

Veenhoven, R (2016d). *Measures of Happiness*, World Database of Happiness, Erasmus University Rotterdam

Available at; http://worlddatabaseofhappiness.eur.nl/hap_quer/hqi_fp.htm