

‘VERY HAPPY’ IS NOT ALWAYS EQUALLY HAPPY

On the meaning of verbal response options in survey questions

Tineke DeJonge¹, Ruut Veenhoven² and Lidia Arends³

Journal of Happiness studies, 2015, 16 (1),77-101, ISSN 1573-7780
DOI 10.1007/s10902-013-9497-9

ABSTRACT

Survey research is based on questioning and respondents typically answer to questions by picking one of several response options. These response options are labeled verbally with terms such as ‘very happy’ or ‘fairly happy’. Response scales differ in the number and wording of response options and this could affect the degree of happiness denoted by such words. If so, scores on differently worded questions on the same topic cannot be compared and this reduces the opportunities for research synthesis greatly. Several methods for homogenization have been developed for dealing with that problem, among which the recently proposed ‘Scale Interval Method’ in which judges rate the interval denoted by verbal response options on a continuous 0 to 10 scale. This method allows a view on the size of the problem. Application to commonly used survey questions on happiness in Dutch language reveals considerable differences, the implications of this for research synthesis are discussed.

Keywords: happiness, satisfaction with life, subjective wellbeing, response scales, scale interval method, research synthesis

1 INTRODUCTION

Survey research is a major method used in the social sciences, in particular in the study of subjective wellbeing. Survey research is largely based on standard questions to which respondents answer by picking one of several response options. Response options are often labelled verbally, in the case of happiness research with terms such as ‘very happy’ or ‘fairly happy’. Such response scales differ in the number and wording of response options as the following examples illustrate.

Different questions on the same topic

A question on happiness in the periodical Dutch Household Survey reads:

To what extent do you consider yourself a happy person?

- *very happy*
- *happy*

¹ J.J. DeJonge, Erasmus University Rotterdam, The Netherlands, Erasmus Happiness Economics Research Organization. e-mail: dejonge@ese.eur.nl

² R. Veenhoven, Erasmus University Rotterdam, The Netherlands, Erasmus Happiness Research Organisation and North-West University South-Africa. e-mail: veenhoven@ese.eur.nl www2.eur.nl/fsw/research/veenhoven

³ L.R. Arends, Erasmus University Rotterdam, The Netherlands, Institute of Psychology & Institute of Pedagogical Sciences, and Erasmus MC, Biostatistics. e-mail: arends@fsw.eur.nl

- *neither happy nor unhappy*
- *unhappy*
- *very unhappy*

The same topic is also measured using questions with a different set of response options, such as this question in the International Social Survey Program:

If you were to consider your life in general, how happy or unhappy would you say you are, on the whole?

- *completely happy*
- *very happy*
- *fairly happy*
- *neither happy nor unhappy*
- *fairly unhappy*
- *very unhappy*
- *completely unhappy*

Both questions offer the option ‘very happy’, but do these options denote the same degree of happiness? Probably not. The difference between ‘very happy’ and the next option is likely to be larger in the first case, ‘happy’, than in the second, ‘fairly happy’. Likewise ‘very happy’ is likely to denote a higher degree of happiness in the first case, where it is presented as the highest option, than in the second case where ‘very happy’ comes after ‘completely happy’.

Questionable comparability

This difference in the phrasing of questions is no problem when surveys are analysed separately, but it limits the comparability of findings gathered in different surveys that used different questions.

The case of happiness research illustrates the scope of this problem. Survey research on happiness took off in the 1970s in the wake of the Social Indicator Revolution. To date about 10,000 empirical studies have been done to assess happiness (Veenhoven 2013), and some 1000 slightly different questions have been used (Veenhoven 2013a). About half of the differences in questions are in the number and wording of response options.

This diversity is treated in two ways: one way is to abstain from any comparison when rating scales are not fully identical. This means that most of the findings on happiness are incomparable and thus lost for synthetic analysis. The other approach is to ignore the problem, typically by assuming that the rank numbers of the response options reflect the degree of happiness denoted and treating these numbers as metric values that can be transformed to the same range. This latter approach carries the danger of producing nonsense.

Research question

This begs the question of how serious the comparability problem really is. If differences are marginal, we can continue in the second way. If not, we must either abstain from comparison at all or develop better methods for scale transformation. Since we cannot address all comparability issues, we limit ourselves to the comparability of responses to identical response options that figure in non-identical scales, such as the option ‘very happy’ in the two questions above. Does ‘very happy’ mean just as much happiness in these cases? More formally formulated our research question reads: To what extent does the meaning attached to identical response options differ when used in the context of non-identical response scales?

Plan of this paper

We will start with a more detailed discussion of the variation in rating scales used for survey questions on happiness in section 2. This gives us insight in the different aspects of the variations and thus in the problem we are talking about. In section 3, we will review some

conventional scale transformation methods and explain why these methods are inadequate to answer our research question. Next in section 4 we introduce a recent method that overcomes these shortcomings. In this method judges are deployed to estimate the degree denoted by each of the verbal response options (Veenhoven 2008). In section 5 we report how we applied this method and we describe our results in section 6. We close with a discussion on the strengths and limitations of this approach in section 7 and the main conclusion in section 8.

This paper is restricted to the preliminary research question of how serious the differences in interpretations really are and does not address the subsequent question what methods are the most suited for correcting such differences. That latter issue is addressed in DeJonge et al (2013).

2 DIFFERENCES IN RESPONSE SCALES

Most people have a positive perception of their own wellbeing, at least in the western world. As a result, the distribution of responses to questions on happiness is skewed, with a long tail on the left that represents 'negative' outcomes (Diener & Diener 1996, Cummins 2003). Irrespective of the scale used, this skewness has to be kept in mind when interpreting the results of such measurements.

Within the large set of existing measures of happiness, the number of response options and the distinction between verbal and numerical response scales are obvious variations. To meet the skewness of the distribution, in the past verbal scales have been devised that are skewed due to mainly positively formulated response options. An example of such a scale is the one used by Statistics Netherlands to measure satisfaction with life, consisting of the response options 'Extraordinarily satisfied', 'Very satisfied', 'Satisfied', 'Fairly satisfied' and 'Not very satisfied'. Only the latter of these responses is formulated negatively. The idea behind this rather asymmetric scale at the time it was devised was that it would give the possibility for more variation in the responses than if a more symmetric scale was used. The satisfaction with life scale of Statistics Netherlands is a unipolar scale: all response options contain the word 'satisfied'. This differs from a bipolar scale, where in the response options, for example, the word 'dissatisfied' would be used as the opponent of 'satisfied'. Furthermore, a scale does not necessarily need to have a neutral midpoint dividing it into a positive and a negative pole and the end points of different scales may vary in the extremity of the wording used, for example 'extraordinarily' is more extreme than 'very' but both are subject to the respondents interpretation of the words and this will vary from respondent to respondent, and each variation will influence the response patterns (Cummins 2000).

Most of the variations discussed above hold for both verbal response scales and for numerical scales. Although numbers are used on a numerical scale to express the respondent's degree of happiness, it is still necessary to use words to describe what the end points of the scale mean, and it is this wording which defines whether the scale is unipolar or bipolar. The wording of such descriptions can include the subject of measurement, as in 'dissatisfied' – 'satisfied' or leave to the respondents how they interpret the end points of the scale when a formulation is given in terms like 'best possible' – 'worst possible'.

Other variations in numerical scales are the visual orientation, which can be vertical or horizontal, and the labelling of the anchor points that can go from negative to positive, for example -5 to +5, consists only of positive numbers and zero starting at 0 or 1, or there can be no numbering (Mazaheri & Theuns 2009). In an experiment done by Schwartz et al (1991) using an 11-point numerical scale with extremes labelled from 'Not at all successful' to 'Extremely successful' and ranging from -5 to +5 only 13 per cent of the respondents gave an answer between -5 and 0. When the range changed from 0 to 10, the percentage of answers at

the lower end of the scale changed to 34 per cent. A similar result was found by Sangster and colleagues (2001). From this experiment Schwarz and his colleagues concluded that a numerical scale starting at zero suggests the absence or presence of the subject under study, which makes the scale unipolar. If conversely, one half of the scale is negative and the other half is positive, then the positive values are related to the presence of the subject one is interested in, whereas the negative values represent the opposite. Schwartz et al also suggest that scales that are intended to assess the intensity of a single attribute, for example happiness, should follow a zero-to-positive-values format to emphasize that the question pertains to the absence or presence of this specific attribute, rather than the presence of its opposite. This suggestion is underpinned in an elaborated discussion on happiness as a variable in Kalmijn (2010). In his thesis, Kalmijn devotes a section to our perspectives on the nature of happiness and satisfaction, the difference between intensity and extensity variables, the polarity of happiness scales and the level of measurement.

The response scale cannot be seen separately from the related lead question. The combination of a scale and a question makes a 'survey item'. The variations in the wording of the questions also lead to numerous different survey items. Furthermore, the time frame a question relates to leads to more variations. For example, the question can refer to satisfaction with life over the life time or just at this moment or in the last four weeks. Moreover, the question can contain a keyword such as the word 'happy' in the question 'Are you happy with your life?', where the subject can either be or not be explicitly formulated in the response options, but likewise be formulated as 'Do you feel ...?' with the keyword only mentioned in the labels of the response options of the related scale.

These are just some examples of the variations in the wording of the questions used in happiness research. Of course there are many more variations one can think of and a comprehensive description of such questions and a discussion on these variations is given in Saris and Gallhofer (2007). Additionally a systematic overview of all the variations in survey items on happiness can be found in the collection 'Measures of Happiness', which is part of the World Database of Happiness (Veenhoven 2013a).

3 CONVENTIONAL SCALE TRANSFORMATION METHODS

Several methods have been developed to transform ratings on different response scales to a common one, typically a 0-10 numerical scale. Some of these scale transformation methods are applied in the World Database of Happiness, in particular in its collection of 'Happiness in Nations' (Veenhoven 2013b). In this section we describe two conventional methods and why they fall short to answer the research question to what extent the meaning attached to identical response options differs when used in the context of non-identical response scales.

3.1 Linear Stretch

One commonly used conventional transformation method is the Linear Stretch Method, which is most applicable for questions that use a numerical response scale. Scales with five or seven response options are typically stretched to obtain a common range from for example 0 to 10. This is done in such a way that the lowest number assigned to a response option is always projected onto 0 and the highest number onto the highest value of the numerical scale, and all the intermediate options are given equally distanced numbers in between: for a 5-point verbal scale the transformation to a 0-10 scale according to this method results in [0.0; 2.5; 5.0; 7.5; 10.0]. The transformed sample mean follows from the conventional frequency approach according to which it is equal to the sum of the transformed values of all response options multiplied by their respective measured relative frequencies. When a verbal scale has to be

transformed in this way, an initial step is to assign numerical values to verbal response options, typically using consecutive numbers, such as 4 for the happiest option on a 4-step scale and 1 for the least happy option.

The Linear Stretch Method has many serious disadvantages. The two most prominent of these disadvantages are one, the assumption made that the distances between the response options are equal, and two, even more problematically, the assumption that the labelling of the response options is irrelevant to the analysis, though not for the respondent. Despite these disadvantages, the Linear Stretch Method is still applied, for example it is used in the World Database of Happiness for numerical scales with at least seven points to transform them to comparable scales with a 0 to 10 range.

The Percentage of Scale Maximum

Another example of where the Linear Stretch Method is applied is in the percentage of scale maximum (%SM) method developed by Cummins (1997, 2003). In this method Likert scale data are transformed to a standard form with a range from 0 to 100. In the %SM-method a score of '0' is given to the lowest scale anchor up to 'n' to represent the highest scale anchor. Any mean score on this scale can subsequently be converted into %SM units by converting the score into a percentage of the scale maximum value as: $\%SM = (\text{mean score}/n) * 100$.

The %SM-method encounters the same disadvantages as mentioned above for the Linear Stretch Method, since it is based on a simple linear relationship between the response options and disregards their labelling. This can be illustrated by a practical example from an e-mail discussion about the discrepancy between the %SM-scores on satisfaction with life as a whole for Australia and New Zealand⁴. In the Australian Unity Wellbeing Index (AUWI) project, survey 21, there was a %SM-score of 78% (Cummins 2009). Applying the %SM-method to the results found in the New Zealand General Social Survey⁵ (NZGSS) 2008, would give a rather lower score of 70%, whereas the discussants would expect the two countries to be virtually identical on such measures. The reason for the discrepancy was found in the differences in scales that were used to compute the results. In the AUWI a 0 to 10 numerical scale was used, whereas the NZGSS used a 4-point verbal scale with response options labelled as (1) dissatisfied/very dissatisfied, (2) neither satisfied nor dissatisfied, (3) satisfied and (4) very satisfied. When the %SM-method was applied these response options were converted to, respectively 0, 33.3, 66.6 and 100, irrespective of their labelling. The remark was made by one of the discussants that 33.3 was perhaps a low score for respondents that state they are neither satisfied nor dissatisfied and that a value of 50 would seem to be fairer.

As a solution it was suggested that the original NZGSS-scale could be considered as a 5-point scale with the lowest two categories grouped and given scores of 100, 75, 50 and 12.5. In this way the middle category would get 50 points and the lowest categories, 0 and 25, would be averaged out. Under this method, the average score would equal 77.0, which would be very close to the Australian average, as might be expected. From this practical example, it becomes very clear that, when comparing the results of different surveys, the labelling of the response options cannot be neglected when converting verbal scale outcomes to a common numerical scale.

⁴ E-mail discussion of December 2009 between participants to the OECD/ISQLS/ISTAT meeting "Measuring subjective well-being: an opportunity for National Statistical Offices?", Florence, 23-24 July 2009

⁵ http://www.stats.govt.nz/browse_for_stats/people_and_communities/Households/nzgss_HOTP2008.aspx, Excel tables for NZGSS 2008 HOTP.xls, Table 10 cont.

3.2 Semantic Judgement of Fixed Word Value

Several attempts have been made in the course of happiness research to develop better methods to cope with the heterogeneity in response scales. What many of these alternative methods have in common is that they make use of expert ratings (Veenhoven 1993, Bălăţescu 2002, Lim 2008) by getting a group of experts to rate the verbal labels of response options on a common numerical scale.

An early example of such a method is that of Jones and Thurstone (1955) who requested approximately 900 respondents to rate 51 verbal qualifications on a 9-point Likert scale separately. A value on a common interval scale and a standard deviation were calculated for each qualification. The result was a list of the 51 qualifications ordered on the basis of their value on the common interval scale. This method, which we have classified as the Semantic Judgement of Fixed Word Value Method, is also applied in the World Database of Happiness to obtain comparable average scores. Veenhoven (1993) and 12 co-workers rated the degree of happiness denoted by the verbal labels of 29 commonly used survey items on a numerical 0 to 10 scale. For example, the label ‘very happy’ was an option in 8 of the 29 items and was given a rating varying from 9.2 to 9.4 resulting in an overall mean of 9.3, whereas for the label ‘not very happy’ an overall mean of 3.7 was found. To this day, these results are used to transform responses in the WDH for scales for which linear stretching falls short.

The Semantic Judgement of Fixed Word Value Method overcomes the disadvantages of presumed equidistance and the neglecting of the labels that are associated with the Linear Stretch Method. The Semantic Judgement of Fixed Word Value Method, however, also has some weak points. Kalmijn (2010) mentions that the fixed values applied in the World Database of Happiness:

- are based on expert judgements that do not necessarily reflect the views of non-expert respondents
- have been rated by Dutch experts on basis of the English version of the questions, thus implicitly assuming that the feelings associated with an item are not affected by its translation from Dutch into English
- do not take into account the phrasing of the lead question, nor the number and the labels of the alternative response options and their position on the scale

4 SEMANTIC JUDGEMENT OF WORD VALUE IN CONTEXT

To counter the shortcomings of the Semantic Judgement of Fixed Word Value method, Veenhoven (2008) developed the Happiness Scale Interval Study, in which ‘judges’ rate the degree of happiness denoted by each of the verbal response options in the context of the full item. The judges are asked to identify the interval on a 0-10 numerical scale that corresponds with a verbal response option such as ‘very happy’ using a web-based Scale Interval Recorder (Veenhoven & Hermus 2006). This method is discussed in detail in (Kalmijn 2010; Kalmijn et al 2011) and we have classified it as the Semantic Judgement of Word Value in Context Method.

The Scale Interval Recorder

A series of survey items is presented on a computer screen to the judges. Items are presented one by one on the left side of the screen and each item presented consists of a question and its corresponding verbal response scale with options given in the judges’ mother tongue. A screen shot of the Scale Interval Recorder from a study presented to Dutch judges is given in Fig. 1. On the right side of the screen a vertical bar scale is displayed with a number of small

horizontal slides on it, the number of which is equal to the number of response options minus one. The judges have to shift the slides until they feel that the intervals on the vertical bar correspond to the meaning of the words as used for the verbal response options. Note, the response options that are displayed next to the bar move simultaneously with the slides to the level of the mid interval value of each interval.

Looking at Fig. 1 it can be seen that the extremes of the numerical bar scale are labelled 'Worst possible' and 'Best possible'. In the terminology of Saris and Gallhofer (2007) these labels are called 'fixed reference points'. What worst and what best means, is left to the interpretation of the judges. The labelling of the extremes is thus semi-abstract which makes them applicable to all questions presented to the judges and independent of the subject of an individual question. An additional advantage of this semi-abstract labelling is that the judgement is not influenced by the extremity of the wording used for the labels of both end points of the continuum.

Since the main aim of the Happiness Scale Interval Study is to improve how we compare happiness across nations, the items included are restricted to those that have been applied in studies of general populations in nations. A complete list of all items ever considered since the start of the study in 2005 can be found on the website of the World Database of Happiness, in the section 'Scale Interval Study'. The direct link to this section is: (http://worlddatabaseofhappiness.eur.nl/scalestudy/scale_fp.htm).

5 HOW CONTEXTUAL DIFFERENCES IN WORD VALUE WERE ASSESSED

The question addressed in this paper was: To what extent does the meaning attached to identical response options differ when used in the context of non-identical response scales? We used the Semantic Judgement of Word Value in Context Method described above to assess such differences in ratings of happiness.

Survey questions

We considered twenty survey questions on happiness in the Dutch language. Eighteen of these questions were taken from past and recent national surveys fielded in The Netherlands. Two of the questions were variations on existing questions. The first of these was the satisfaction with life item used by Statistics Netherlands and described in section 2 of this paper with the phrasing 'satisfaction with life' in the lead question and the word 'satisfied' in the labels of the response options replaced by the word 'happy'. The second variation was similar to the first one with the subject happiness of an existing item replaced by the subject satisfaction with life in both the leading question and the labels of the response options. These experimental items were introduced to allow respondents interpretations of happiness scales to be compared with their interpretations of satisfaction with life scales. A complete overview of all the twenty survey questions used in this research can be found in the 'study list' on the website of the Happiness Scale Interval Study, mentioned above. The studies Dutch6 and Dutch7, each of which comprises ten questions, should be selected.

An overview of the original Dutch wording of the selected items for this paper and their translation into English, ordered according to their number of response options and preceded by the code they have been assigned in the collection 'Measures of Happiness' of the World Database of Happiness, is given in appendix A.

Judges

Both the Dutch6 and Dutch7 studies were done using students from the Erasmus University Rotterdam and employees of Statistics Netherlands and The Netherlands Institute for Social Research. The employees could participate in both studies. As a result the number of participants in each study was rather large, amounting to 392 judges for Dutch6 and 359 for Dutch7.

University students are widely employed as research participants in social science research, but it is doubtful whether they are representative of the general population. Cummins (2003) argues they are not because they are on average younger, better educated, to tertiary level and come from more privileged backgrounds than most people in their respective populations. From this he concludes that data derived from students are likely to be both biased, in respect of general population data, and more homogeneous. He states that if this proposition is accepted, then analyses based on student data should be as cautiously interpreted as data from any other non-representative group. Most of the arguments Cummins brings forward for students could also be applied to the group of employees used for this research. On average they were about fifteen years older than the students, but like the students the majority of the employees invited to participate in the study had been educated to a tertiary level. In addition, many of the employees were very familiar with survey research, which probably made them even less representative of the general population than the students.

The Scale Interval Recorder though, is an instrument to assess how people interpret words in common language. We believe that the appreciation students have of their own happiness may on average differ from that of the general population, but that they do not differ in the meaning they assign to words in common language. There were however, some small differences between students and employees in the mean values they assigned to the boundaries between response options, as can be seen in the table in appendix B. Despite these differences the conclusions that can be drawn from the judgements of both groups are equivalent. This gives confidence that the outcomes for the general population will not go in a totally different direction and that the conclusions based on the above study will be valid for the general population. In the remainder of this paper only the combined results for both students and employees are used. A division of the results into Dutch6 and Dutch7 was not relevant for the analyses, even though the group of judges was not entirely the same. The initial division into two studies was mainly done to prevent there being too many questions to judge within one study and Dutch6 and Dutch7 were seen to be essentially part of the same study for the purpose of the analyses.

6 RESULTS

The aim of the above mentioned Scale Interval Study is to transform responses on different verbal scales to a common numerical scale. In this paper we use the technique to answer the preliminary question of whether such transformation is required. The research question was: To what extent does the meaning attached to identical response options differ when used in the context of non-identical response scales?

6.1 The keywords ‘happiness’ and ‘satisfaction with life’ and the degree of appreciation

We considered two kinds of questions on subjective appreciation of life: questions that use the keyword ‘happiness’ and questions that used ‘life satisfaction’ as the keyword. These are not the only terms used in questions on this topic, but data on other keywords fall short. In this

analysis we did not consider possible differences in meaning attached to these terms, but focused on differences in ratings of degree of appreciation in each separately.

From Fig. 2 it becomes clear that the Dutch speaking judges who participated in our study assigned almost identical degrees of appreciation to the response options of equivalent response scales for happiness and satisfaction with life⁶. Only for the second pair of items did the boundaries between the response options not fully coincide but the difference in degree of appreciation for this pair was still not noteworthy. This difference could be designated as negligible when compared to differences in degree of appreciation of similar response options in dissimilar scales, for example the response option ‘fairly satisfied’ in the 4-point scale of the first pair was appreciated totally different by than for the same option in the 5-point scale of the second pair.

The findings from Fig.2 do not imply that happiness and satisfaction with life are the same construct. Despite the equivalence in degrees of appreciation assigned to response options, respondents may rate their feeling of ‘happiness’ somewhat differently from how they appreciate their ‘satisfaction with life’. The equivalence in degrees of appreciation merely means that a comparison of ratings for happiness and satisfaction with life is not disturbed by a difference in the degrees of appreciation assigned to the response options in equivalent scales. We also need to remark that the similarity in degrees of appreciation by Dutch judges does not mean that this will be the case in general. In other languages and cultures this may be different. In this paper we will not go further into this, but will attend to this topic in a paper on the equivalence of rating scales using different keywords. Given the results for the Dutch speaking judges, we feel it is justified not to distinct between questions that use the term ‘happiness’ or those that use ‘satisfaction with life’ for the results presented in this paper.

6.2 The meaning of ‘happy’ and ‘satisfied’ in the context of the response scale

We started with the question ‘Are you happy with your life?’, with response options ‘yes’ and ‘no’. The interpretation of these options and that of an item with four response options are presented in Fig. 3.

If only the options ‘yes’ and ‘no’ are offered, the continuum is not partitioned into two intervals of equal length, the split is made at the value 5.8. This may partly be influenced by the fact that in the Dutch school system tests are graded on a numerical scale from 1 to 10, where a grade below 5.5 mean that you have failed the test.

The 4-point scale item in Fig. 3 is illustrative for the difficulty of comparing survey results for different items. The words ‘yes’ and ‘no’ in the 2-point scale item can be replaced by ‘satisfied’ and ‘not satisfied’, however, these qualifications of the degree of happiness do not return in the 4-point scale. In this latter scale the wording chosen to express the degree of satisfaction is stronger or weaker than ‘satisfied’ and ‘not satisfied’. The most notable of the results for the 4-point scale is that, although the wording chosen for the most negative option cannot be superseded by something that is more extreme, the average length of the interval it has been assigned by the judges is rather large. The reason could be that the preceding response option is formulated rather moderately.

More insight into the effect the wording used for labelling response options has on how a scale is interpreted can be gained by comparing the results for the three differently labelled 3-point scales shown in Fig. 4.

⁶ In the first pair, the English wording used for the one best option are for the item on happiness not fully comparable to that for the item on satisfaction with life. The difference is in the words ‘fairly’ and ‘quite’. In the Dutch version, however, the word ‘tamelijk’ is used for both items.

For the first two items in Fig. 4, which have equally labelled extremes, the wording chosen for the label of the middle option can be seen to be crucial for the interpretation of the scale. A negative formulation comes at the cost of the lower extreme and a positive formulation at the cost of the upper extreme. A more extremely labelled end point, as in the third item, turns out to reduce the valuation of the word 'satisfied' to a much lower position on the continuum. In this scale 'satisfied' no longer acts as the upper part of the continuum, the lower bound of this option has dropped below 5. The percentages of happy people measured with a response option labelled 'happy' are clearly shown in Fig. 4, but it is obvious that the different scales cannot be compared in a straightforward manner. The meaning of the word 'happy' shows a large contextual dependency on the composition of the response scale.

This becomes even more obvious from Fig. 5 in which two items with asymmetric response scales, one on satisfaction with life and one on happiness, are presented next to symmetric variants. Comparing the two items on satisfaction with life, it is striking to see that the response option 'fairly satisfied' in the asymmetric scale is positioned in the lower part of the continuum, whereas in the symmetric scale this is a degree of satisfaction with life on the other, non-overlapping part of the spectrum. In addition it is remarkable that, although the asymmetric scale consists of five response options, the option 'not very satisfied' covers more than one third of the continuum, probably due to the absence of a truly negatively labelled option.

Looking at the items on happiness in Fig. 5, it can be noted that the upper part of the scales is equal for both items. The interpretation of this scale shows only minor differences between the two items. The distortion comes in the lower part of the scale for the third item which is not symmetric compared to the upper part. The word 'unhappy' in the asymmetric scale seems to denote something different than the same word in the symmetric scale. Furthermore it is noteworthy, from Fig. 5, that the last three items all have a neutral response option in the middle. This neutral option is interpreted as positioned in the middle of the continuum, where the length of the interval seems to depend on the wording of the surrounding response options. A weaker labelling of the directly neighbouring response options seems to trim down the length of the interval of the neutral option.

6.3 The effect of the wording chosen for the extremes of the response scale

The wordings chosen for the extremes of the response scales are especially interesting, the extremes of the numerical bar scale on which judges have to rate the verbal response options are labelled 'Worst possible' and 'Best possible'. This scale is all inclusive offering space to every degree of satisfaction with life or happiness one can think of, it can even be combined with both questions on happiness and satisfaction with life, since the topic is not explicitly part of the wording of the extremes. These characteristics are not obvious for a response scale, as demonstrated by the examples given above. Most often the extremes are expressed in terms of the topic they relate to and, although they may differ in the intensity they express, they do not always make a scale all inclusive.

The labelling of the extremes of a response scale is depicted in Fig. 6. The two items with a symmetric scale depicted in Fig. 5, return in Fig. 6 together with two items that also have a symmetric scale, but where the word 'completely' is used instead of 'very' for the labelling of the extremes. Comparison of the first two items shows that the intervals for the response options labelled with a 'completely' are smaller than when they are labelled with a 'very'. This can be attributed to the fact that more than completely is not possible, whereas one can be 'very satisfied' but still not be 'completely satisfied'. The word 'completely' as used for the items in Fig. 5 make the scales that they belong to all inclusive with respect to the topic they refer to.

Comparing the extremes of the third and fourth item in Fig. 6, it is clear that the word ‘completely’ does not replace the word ‘very’ in the scale of the fourth item, but it has been introduced to label an extra response option on both sides. As a result this last item offers respondents the largest choice of options. The intervals assigned to these options however, seem to be somewhat compressed compared to the 5-point scales. The space reserved for the extremes of this 7-point scale is rather limited compared to the other response options. One could question whether a 7-point verbal scale has an added value over a 5-point verbal scale. Response options formulated with wordings such as ‘completely’ or ‘not at all’ may tempt judges to assign a zero-width interval to them by choosing the upper and lower bound both equal to one of the extreme values of the continuum (Kalmijn 2010). These labels all express some kind of limit that cannot be exceeded however, for the selected items with less than five response options, these zero-width intervals occur very rarely. For the items with at least five response options the percentage of judges that assigned a zero-width to the extremes is displayed in Fig. 7.

Although the formulation ‘extraordinarily’ is rather extreme, strikingly it does not lure judges to assign a zero-width interval to one of the extremes of the verbal scale. Yet the percentage of zero-width intervals for the first item in Fig. 4 differs from that for the other 5-point items in that the largest percentage belongs to the upper extreme of the scale. Maybe the percentage of zero-width judgements for the lower response option of the second item is more important. The percentage is not very high, but the response option it applies to, is not extreme. The zero-width interval for this item might be attributed to the fact that the range from ‘very happy’ to ‘not very happy’⁷ can be considered to be a complete, though not all inclusive scale making the response option labelled ‘unhappy’ an outside class. With respect to this, the composition of the scale belonging to the third item, starting with ‘very unhappy’ and ending with ‘very happy’ seems to be more logical. Most notable however, are the percentages for the two items with the words ‘completely’ in the labelling of their response options. There are no superlatives for the response options with these labels. For the 7-point item the zero-width percentages make it plausible that the formulation of the extremes for this scale have no added value. From this it may be concluded either that a 5-point scale would do, what would amount to a scale as for the fourth item, or that a different and a less extreme wording should be chosen for the response options labelled ‘very unhappy’ and ‘very happy’ in the 7-point scale under the condition that a verbal response scale would be preferred over a numerical response scale.

6.4 The effect of the number and wording of response options on the central tendency

Although the outcomes of the study as presented in the previous sections are probably not very surprising, and as one could have expected, it is interesting to see what the effect of these differences in the number and wording of response options has on the central tendency of the scores on happiness and satisfaction with life. A glimpse on this can be found in the scores on satisfaction with life and happiness taken from the Permanent Survey on Living Conditions of Statistics Netherlands⁸. In this survey the scores are based on the first and third items presented in Fig. 5.

The response scale for the item on satisfaction with life is an exceptional case, because it is asymmetric with only one negative response option and options formulated in extreme

⁷ It must be noted here, that ‘not so happy’ is somewhat closer to the original Dutch phrasing for this response option. The word ‘very’ instead of ‘so’ is chosen however, for the translation for reasons of comparison with other scales, but this does not change the conclusion.

⁸ <http://statline.cbs.nl/StatWeb/publication/?DM=SLEN&PA=60027ENG&D1=43-52&D2=0&D3=a&LA=EN&HDR=T&STB=G1,G2&VW=T>

terms at the positive end of the response scale. Due to this, the interpretation of the response options 'satisfied' and 'fairly satisfied' seems to be forced to end in a too low part of the numerical scale. The response option 'fairly satisfied' is judged to be less positive than the neutral options of the other items shown in Fig. 5. The scale of the item on happiness is more or less symmetric, with a neutral option in the middle, but with the 'extreme' for the lower end labelled 'unhappy' making it an outside class as discussed in section 6.2.

Over the years the scores for these items in the Netherlands have been stable but they show a remarkable difference in the central tendency of the response on happiness and satisfaction with life. Two thirds of the population claims it is 'happy' whereas less than half of the population believes it is 'satisfied'. Conversely, over forty per cent of the Dutch seems to be 'very satisfied' or even 'extraordinarily satisfied', which is about twice as much as the share of people indicating they are 'very happy'. It is very likely that the answer given by a respondent is influenced by the position of the response option on the scale. Someone who is satisfied with his or her life would consider a response option, even if it is labelled 'satisfied', in the middle of the scale not in concordance with this, this could tempt these persons to choose the option 'very satisfied', since this is the position on the scale that is more in harmony with their perception of satisfaction with life.

In practice this difference in outcomes for happiness and satisfaction with life is often overcome by dichotomizing them and for this purpose the population that scored 'very happy' or 'happy' on the question about happiness is declared to be happy. For satisfaction with life the scores on the response options 'extraordinary satisfied', 'very satisfied' and 'satisfied' are combined in a category of people that are assumed to be satisfied with their life. An example of this is given in DeJonge et al (2009). As a result of the dichotomization, the share of happy people nearly equals the share of satisfied people. Although this brings us to the observed difference, one could question whether dichotomization is really justified and if doing so causes us to lose a lot of information.

7 DISCUSSION

A comparison of survey items with as few as 3-point on the scales is already sufficient to see that the number and wording of the response options does matter. If the extremes of a 3-point scale are labelled in a purely oppositional manner, than it very much depends on the wording chosen to label the response option in the middle where the boundaries between this and the two extremes are positioned. A negative formulation comes at the cost of the lower extreme and a positive formulation will trim the length of the upper extreme. In a symmetric scale, independent of the number of response options, a neutrally labelled middle option will be positioned in the centre of the scale. The length of the interval assigned to it will depend on the intensity expressed by the wording used to label the surrounding response options. If the wording is moderate, like in 'fairly happy', the length of the interval will be smaller, than when a more pronounced formulation is applied such as 'happy' or 'very happy'.

Scales that have been deliberately devised to be asymmetric to bring variation in the responses turn out to have an unintended side-effect. These scales may tempt respondents to choose a response option that is positioned on the scale in accordance with their perception even if this is not fully in conformity with the label attached to it. As a result the central tendency of the measurements reflects the asymmetry of the scale instead of the subjective well-being of the respondents.

Sometimes the wording chosen for the labels of the extremes on a scale invite judges to assign a zero-width interval to them, by choosing upper and lower bound equal to one of the extreme values of the continuum. This holds especially for wording such as 'completely'

for which there are no superlatives. If the adjacent response options also have extremely formulated labels, respondents are encouraged even more to assign a zero-width interval to the extremes of the scale, which would reduce their added value. From this it might be concluded that, for verbal scales, extreme options labelled using the word ‘completely’ are more or less redundant if the adjacent options are also expressed in extreme terms, however, to make a numerical scale all inclusive it is preferable to label the anchor points ‘completely’ instead of ‘very’ and even better use the labels ‘worst possible’ and ‘best possible’.

We assumed for the presentation of the results, that it makes no difference for the position of the boundaries whether the word happy or the word satisfied is used to label a response option in an otherwise similar response scale, however, is this assumption true. We intend to investigate this assumption and determine whether it is justified or not.

In the Happiness Scale Interval Study it is assumed that there is no conflict between the personal perception of happiness of a judge and his or her assessment of the response scales (Kalmijn 2010). According to Kalmijn, the only justification for retaining this assumption is that it has never been investigated. In addition to this, it is worth noting that, even if the assumption is true, a response scale does not necessarily offer response options that meet the perception of respondents well, they may force them to choose between two less than optimal alternatives. The least inappropriate option may be ranked in a counterintuitive position by a respondent in between the other response options. As a consequence, the boundaries derived from the assessments by judges may not correspond with how the response options are selected in practice by respondents. To illustrate this, take the item with four response options that we presented in Fig. 3, that offers the response options ‘fairly satisfied’ and ‘very satisfied’. Respondents who are satisfied with their life have to choose between an option that either underestimates or overestimates their perception of satisfaction with life.

We recruited students and employees as judges for our study. We recall from section 5 that there were some small differences between these groups in the mean values they assign to the boundaries between response options, but that the conclusions that could be drawn from the judgements of both groups were equivalent. To ascertain that the outcomes are fully valid for the general population a Happiness Scale Interval Study is required with a group of judges that represents the general population, although this may be difficult to organise.

The study presented in this paper does not lead to an answer as to which scale is the best to use. This however, is not what the study was aimed at. For example from the study by Schwarz et al (1991) it is already obvious that the choice of numbering of closed-ended numerical scales and the labelling of their anchor points affects the distribution to survey answers. What the Scale Interval Study contributes in addition to Schwartz’s study, is that it focuses on verbal response scales and provides a systematic way to express the degree of appreciation denoted by each response options on a continuum from 0 to 10. What our study very clearly revealed is that the degree of happiness denoted by verbal response options, such as ‘happy’ or ‘unhappy’, is strongly affected by both the number of options presented and the wording of these options. Hence findings on the same topic obtained using different verbal response scales cannot be compared well.

In the study presented in this paper, we only looked at survey items taken from Dutch surveys. Since it came into existence however, the international Happiness Scale Interval Study (HSIS) has been conducted in a number of countries and we remind the reader that a complete list of all items ever considered since the start of the study in 2005 can be found on the website of the WDH, in the section ‘Scale Interval Study’. The direct link to this section is: (http://worlddatabaseofhappiness.eur.nl/scalestudy/scale_fp.htm). This fits to the main aim of the HSIS which is to improve the comparison of happiness across nations. The findings from all these studies are very useful to study the influence of culture and language on the

interpretation of verbal response scales. In other languages and cultures the results may be different from what we found for The Netherlands which we will show in another paper.

8 CONCLUSION

The degree of happiness denoted by verbal response options, such as 'happy' or 'unhappy' is strongly affected by the construction of the scale, which is among others reflected in the number of options presented, the wording of these options and whether the scales are unipolar or bipolar. Hence findings on the same topic obtained using different response scales cannot be compared. Conventional methods for scale transformation fail to overcome the differences in degree of happiness or satisfaction with life denoted by the different response options used in different questionnaires. More advanced scale transformation methods are needed before the findings of such studies can be used effectively for research synthesis.

REFERENCES


- Bălătescu, S. (2002).
Problems of transforming scales of life satisfaction
Euromodule workshop Berlin
- Cummins, R. A. (1997).
The Comprehensive Quality of Life Scale – Intellectual/Cognitive Disability (ComQol-I5),
5th edition, School of Psychology,
Deakin University, Melbourne
- Cummins, R.A. & Gullone, E. (2000).
Why we should not use 5-point Likert scales: The case for subjective quality of life measurement.
Proceedings, Second International Conference on Quality of Life in Cities_(pp.74-93).
Singapore: National University of Singapore
- Cummins, R.A. (2003).
Normative life satisfaction: measurement issues and homeostatic model.
Social Indicators Research 64: 225-240.
- Cummins, R. A. (2009).
Australian Unity Wellbeing Index, Survey 21
Report 21.0, May 2009, Australian Centre on Quality of Life, Deakin University, Figure 2.12
- DeJonge, T., Hupkens, C. & Bruggink, J.W. (2009).
Living a happy, healthy and satisfying life. Background paper for the 3rd World Conference of the OECD in Busan, South Korea, Available at
<http://www.oecd.org/dataoecd/63/3/43705841.pdf?contentId=43705842>.
- DeJonge, T., Veenhoven, R. & Arends, L.R. (2013).
Homogenizing responses to different survey questions on the same topic. Proposal of a Scale Homogenization Method using a Reference Distribution
Social Indicators Research, Published online: 24 May 2013, Springer, DOI: 10.1007/s11205-013-0355-6
- Diener, E. & Diener, C. (1996).
Most people are happy
Psychological Science, 7, 181-185
- Jones, L.V. & Thurstone, L.L. (1955).
The Psychophysics of Semantics. An Experimental Investigation
The Journal of Applied Psychology 39 (1), 31-36.
- Kalmijn, W.M. (2010).
Quantification of Happiness Inequality
PhD-thesis, Erasmus University Rotterdam, The Netherlands: Ipskamp Drukkers, Enschede.
Available at <http://repub.eur.nl/res/pub/21777/>

- Kalmijn, W.M., Arends, L.R. & Veenhoven, R. (2011).
Happiness Scale Interval Study, Methodological Considerations.
Social Indicators Research, 102(3), 497-515, DOI: 10.1007/s11205-010-9688-2
- Lim, H.E. (2008)
The Use of Different Happiness Rating Scales: Bias and Comparison Problem?
Social Indicators Research 87 (pp. 259-267), Springer, DOI: 10.1007/s11205-007-9171-x
- Mazaheri, M. & Theuns, P. (2009)
Effects of Varying Response Formats on Self-ratings of Life-Satisfaction.
Social Indicators Research 90 (pp. 381-395).DOI 10.1007/s11205-008-9263-2
- Sangster, R.L., Willits, F.K., Saltiel, J., Lorenz, F.O., & Rockwood, T.H. (2001)
The effect of numerical labels on response scales,
Article presented at the Annual Meeting of the American Statistical Association, Atlanta, GA,
<http://www.bls.gov/osmr/pdf/st010120.pdf>
- Saris, W.E. & Gallhofer, I.N. (2007).
Design, evaluation, and analysis of questionnaires for survey research
Publisher Hoboken, New York, USA, Wiley-Interscience, Wiley series in survey
methodology, ISBN 978-0-470-11495-7, e-ISBN 978-0-470-16519-5
- Schwarz, N., Knauper, B., Hippler, H.J., Noelle-Neumann, E. & Clark, W. (1991).
Rating Scales: Numeric Values May Change the Meaning of Scale Labels.
The Public Opinion Quarterly, 55: 570-582. <http://www.jstor.org/stable/2749407>
- Veenhoven, R. (1993).
Happiness in nations, subjective appreciation of life in 56 nations, 1946-1992.
Studies in Social-Cultural Transformation, No. 2, RISBO, Erasmus University Rotterdam,
Netherlands
- Veenhoven, R. & Hermus, P. (2006).
*Scale Interval Recorder. Tool for Assessing Relative Weights of Verbal Response Options on
Survey Questions*, Web survey program. Erasmus University Rotterdam, Department of
Social Sciences & Risbo Contract Research, The Netherlands
- Veenhoven, R. (2008).
The International Scale Interval Study.
In V. Møller & D. Huschka (Eds), Quality of Life in the new millennium: 'Advances in
quality-of-life studies, theory and research', Part 2: Refining concepts and measurement to
assess cross-cultural quality-of-life (pp. 45-58). Social Indicator Research Series, vol. 35,
Dordrecht, The Netherlands: Springer Press
- Veenhoven, R (2013)
World Database of Happiness: Archive of research findings on subjective enjoyment of life
Erasmus University Rotterdam, Netherlands
Available at: <http://worlddatabaseofhappiness.eur.nl>

Veenhoven, R. (2013a)
Measures of Happiness
World Database of Happiness


Veenhoven, R. (2013b)
Happiness in Nations
World Database of Happiness

Figure 1
Screenshot of the Scale Interval Recorder



HOW HAPPY IS VERY HAPPY?

Happiness Scale Interval Study
version Dutch6



World Database of Happiness
Page 12 of 14

Erasmus University Rotterdam

Question used in survey studies in your country

In welke mate vindt u zichzelf een gelukkig mens?

- erg gelukkig
- gelukkig
- niet gelukkig, niet ongelukkig
- niet zo gelukkig
- ongelukkig

What intervals on the scale fit the meaning of the words used for response options?

best possible
10

0
worst possible

Case 9

Shift the separation lines until you feel that the intervals correspond with the degree of happiness denoted by the words on the right.

Next page

Figure 2
Comparison of the interpretation of verbal response scales by Dutch judges

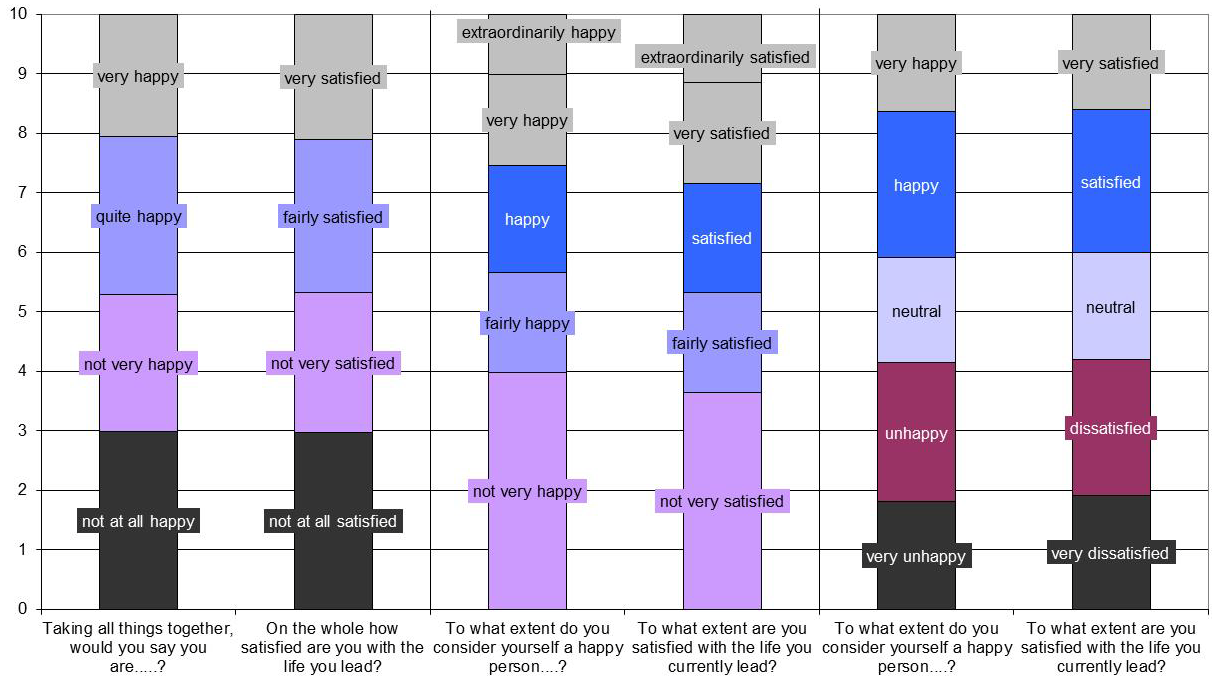


Figure 3
Numerical interpretation of verbal scales on happiness and life satisfaction

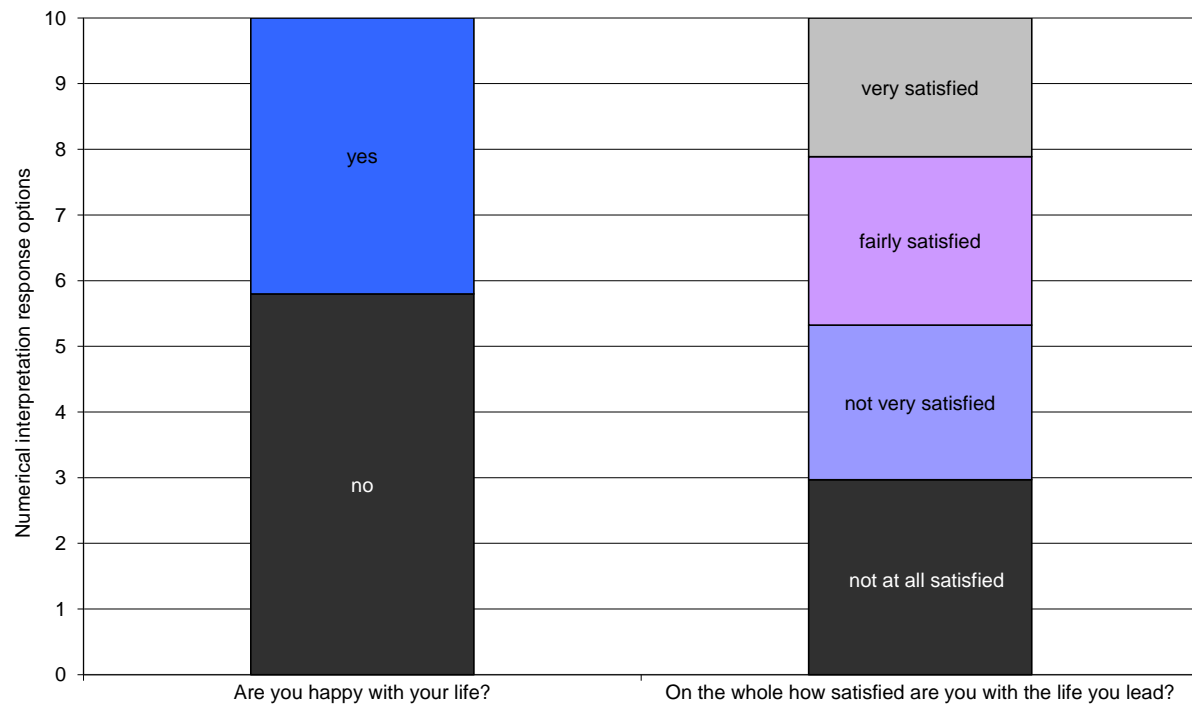


Figure 4
Numerical interpretation of verbal scales on happiness and life satisfaction
(continued 1)

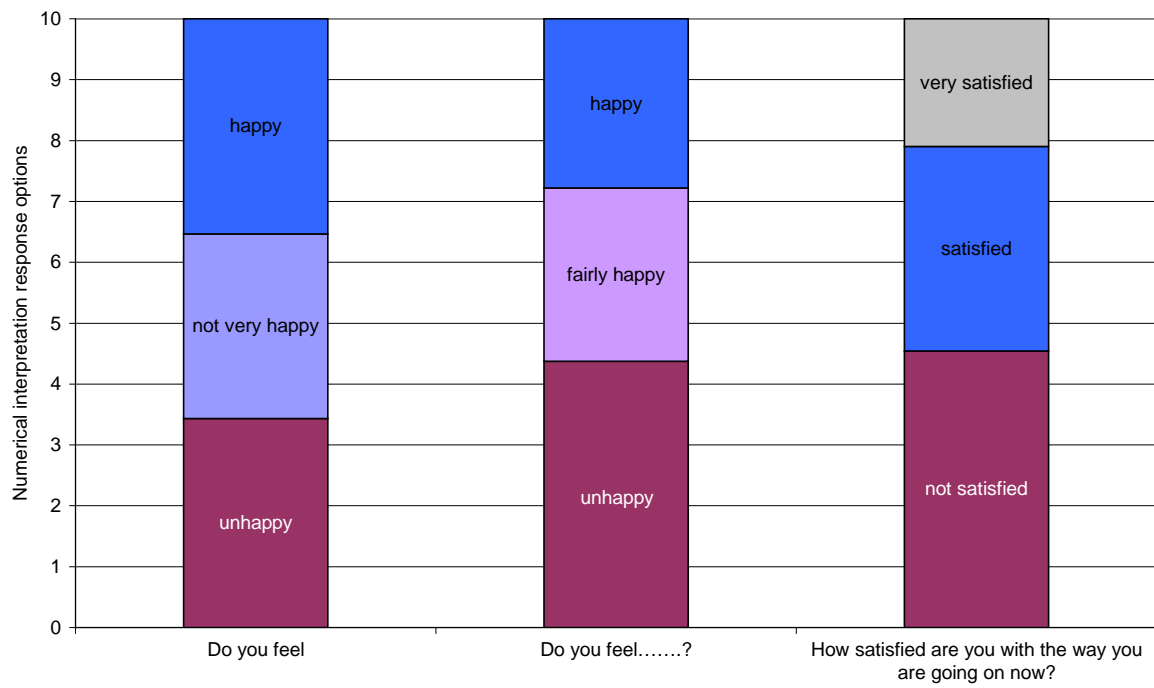


Figure 5
Numerical interpretation of verbal scales on happiness and life satisfaction
(continued 2)

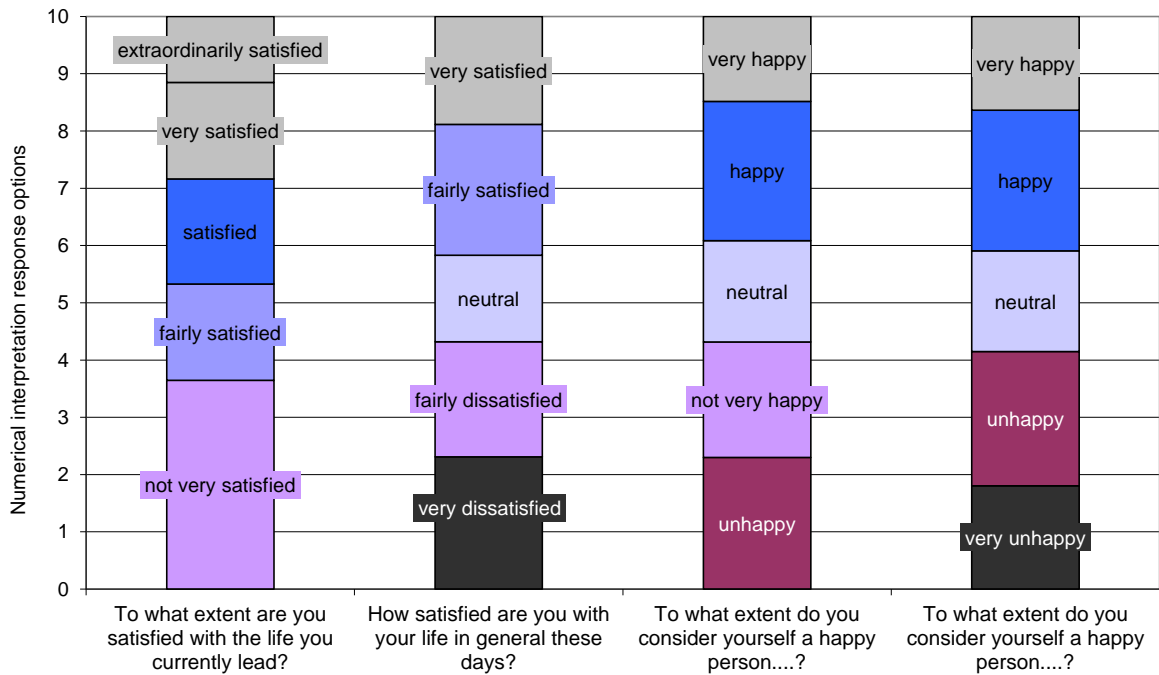


Figure 6
The effect of labelling the extremes on the interpretation of response options

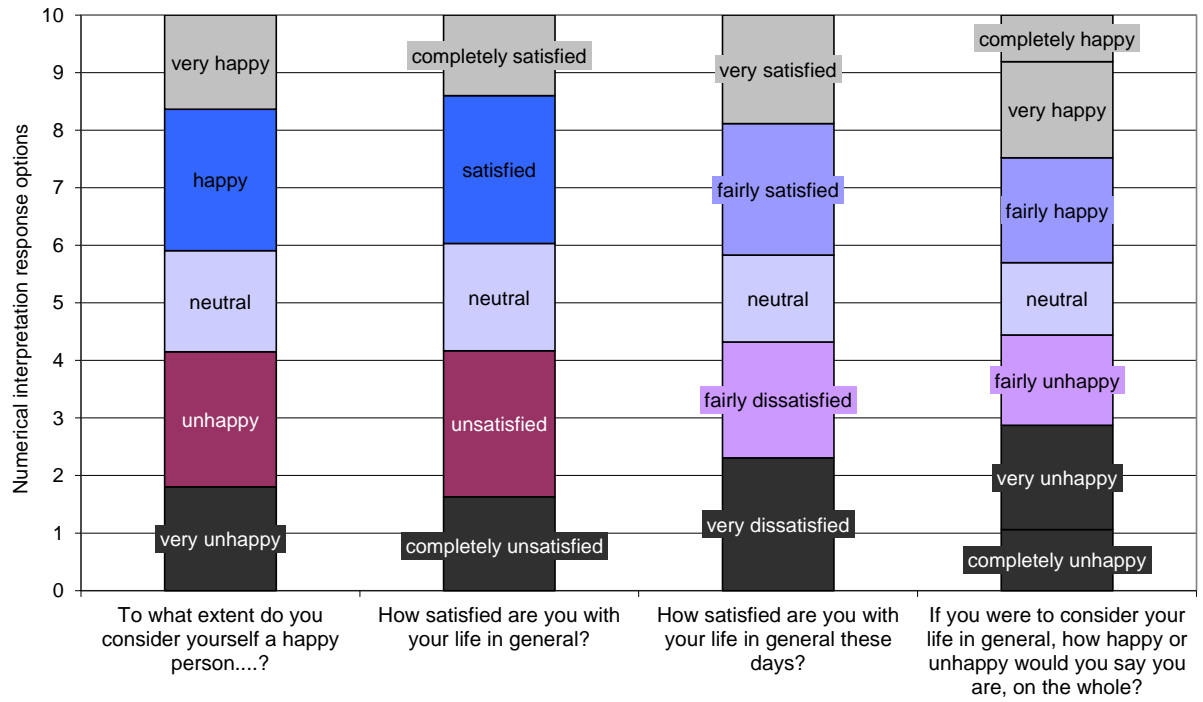
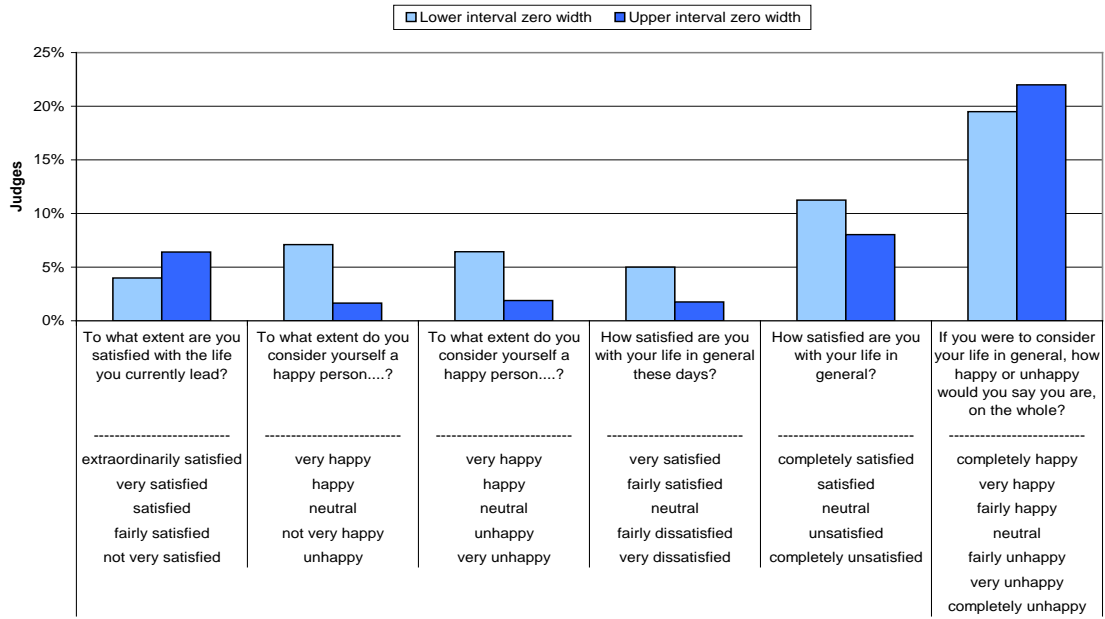


Figure 7
Percentage of extreme response options to which a zero-width interval has been assigned



APPENDIX A**Survey questions on happiness used in this study**

The English translation is also given in this table, but has of course not been presented to the Dutch judges.

Wording in Dutch	Wording in English	Question code
Vindt u zichzelf gelukkig? - ja - nee	Are you happy with your life? - yes - no	O-HL/u/sq/v/2/a
Voelt u zichzelf.. - gelukkig - tamelijk gelukkig - ongelukkig	Do you feel.....? - happy - fairly happy - unhappy	M-FH/u/sq/v/3/d
Voelt u zich - gelukkig - niet zo gelukkig - ongelukkig	Do you feel.....? - happy - not very happy - unhappy	M-FH/u/sq/v/3/a
Nu een vraag over de manier waarop u op het ogenblik vooruit komt: bent u daarover - zeer tevreden - tevreden - ontevreden	How satisfied are you with the way you are getting on now? - very satisfied - satisfied - not satisfied	O-SLS/c/sq/v/3/ab
Hoe tevreden bent u over het algemeen met het leven dat u leidt? - zeer tevreden - tamelijk gelukkig - niet zo tevreden - helemaal niet tevreden	On the whole how satisfied are you with the life you lead? - very satisfied - fairly satisfied - not very satisfied - not at all satisfied	O-SLL/u/sq/v/4/b
In welke mate bent u tevreden met het leven dat u op dit moment leidt? - buitengewoon tevreden - zeer tevreden - tevreden - tamelijk tevreden - niet zo tevreden	To what extent are you satisfied with the life you currently lead? - extraordinarily satisfied - very satisfied - satisfied - fairly satisfied - not very satisfied	O-SLL/c/sq/v/5/d

<p>In welke mate vindt u zichzelf een gelukkig mens?</p> <ul style="list-style-type: none"> - erg gelukkig - gelukkig - niet gelukkig, niet ongelukkig - niet zo gelukkig - ongelukkig 	<p>To what extent do you consider yourself a happy person....?</p> <ul style="list-style-type: none"> - very happy - happy - neither happy nor unhappy - not very happy - unhappy 	O-HP/u/sq/v/5/a
<p>In welke mate vindt u zichzelf een gelukkig mens?</p> <ul style="list-style-type: none"> - erg gelukkig - gelukkig - niet gelukkig, niet ongelukkig - ongelukkig - erg ongelukkig 	<p>To what extent do you consider yourself a happy person....?</p> <ul style="list-style-type: none"> - very happy - happy - neither happy nor unhappy - unhappy - very unhappy 	O-HP/u/sq/v/5/d
<p>Hoe tevreden bent u met uw leven in het algemeen op dit moment?</p> <ul style="list-style-type: none"> - zeer tevreden - tamelijk tevreden - tevreden noch ontevreden - tamelijk ontevreden - zeer ontevreden 	<p>How satisfied are you with your life in general these days?</p> <ul style="list-style-type: none"> - very satisfied - fairly satisfied - satisfied nor dissatisfied - fairly dissatisfied - very dissatisfied 	O-SLW/c/sq/v/5/fb
<p>Hoe tevreden bent u met uw leven in het algemeen?</p> <ul style="list-style-type: none"> - volledig tevreden - tevreden - niet tevreden, niet ontevreden - ontevreden - volledig ontevreden 	<p>How satisfied are you with your life in general?</p> <ul style="list-style-type: none"> - completely satisfied - satisfied - unsatisfied nor satisfied - unsatisfied - completely unsatisfied 	O-SLu/g/sq/v/5/c
<p>Hoe gelukkig of ongelukkig bent u met uw leven in het algemeen?</p> <ul style="list-style-type: none"> - volkomen gelukkig - zeer gelukkig - tamelijk gelukkig - noch gelukkig, noch ongelukkig - tamelijk ongelukkig - zeer ongelukkig - volkomen ongelukkig 	<p>If you were to consider your life in general, how happy or unhappy would you say you are, on the whole?</p> <ul style="list-style-type: none"> - completely happy - very happy - fairly happy - neither happy nor unhappy - fairly unhappy - very unhappy - completely unhappy 	O-HL/g/sq/v/7/a

APPENDIX B Mean and standard error of assessment upper boundaries by employees, students and total

Question code	Wording in English	Mean (employees)	Standard error (employees)	Mean (students)	Standard error (students)	Mean (total)	Standard error (total)
O-HL/u/sq/v/2/a	Are you happy with your life? - yes - no	10.0 5.8	0.1	10.0 5.8	0.1	10.0 5.8	0.1
M-FH/u/sq/v/3/d	Do you feel.....? - happy - fairly happy - unhappy	10.0 7.4 4.7	0.0 0.1	10.0 7.1 4.0	0.1 0.1	10.0 7.2 4.4	0.1 0.1
M-FH/u/sq/v/3/a	Do you feel.....? - happy - not very happy - unhappy	10.0 6.6 3.8	0.1 0.1	10.0 6.4 3.1	0.1 0.1	6.5 3.4	0.1 0.1
O-SLS/c/sq/v/3/ab	How satisfied are you with the way you are getting on now? - very satisfied - satisfied - not satisfied	10.0 8.0 4.8	0.1 0.1	10.0 7.9 4.3	0.1 0.1	10.0 7.9 4.5	0.0 0.1
O-SLL/u/sq/v/4/b	On the whole how satisfied are you with the life you lead? - very satisfied - fairly satisfied - not very satisfied - not at all satisfied	10.0 7.9 5.5 3.1	0.1 0.1 0.1	10.0 7.8 5.1 2.9	0.1 0.1 0.1	10.0 7.9 5.3 3.0	0.0 0.1 0.1
O-SLL/c/sq/v/5/d	To what extent are you satisfied with the life you currently lead? - extraordinarily satisfied - very satisfied - satisfied - fairly satisfied - not very satisfied	10.0 8.8 7.2 5.5 3.8	0.1 0.1 0.1 0.1	10.0 8.9 7.1 5.2 3.5	0.0 0.1 0.1 0.1	10.0 8.8 7.2 5.3 3.6	0.0 0.1 0.1 0.1

O-HP/u/sq/v/5/a	To what extent do you consider yourself a happy person....? - very happy - happy - neither happy nor unhappy - not very happy - unhappy	10.0 8.5 6.3 4.5 2.5	0.0 0.1 0.1 0.1	10.0 8.5 5.9 4.1 2.1	0.0 0.1 0.1 0.1	10.0 8.5 6.1 4.3 2.3	0.0 0.1 0.1 0.1
O-HP/u/sq/v/5/d	To what extent do you consider yourself a happy person....? - very happy - happy - neither happy nor unhappy - unhappy - very unhappy	10.0 8.5 6.1 4.3 2.0	0.0 0.1 0.1 0.1 0.1	10.0 8.3 5.8 4.0 1.6	0.1 0.1 0.1 0.1 0.1	10.0 8.4 5.9 4.1 1.8	0.0 0.1 0.1 0.1 0.1
O-SLW/c/sq/v/5/fb	How satisfied are you with your life in general these days? - very satisfied - fairly satisfied - satisfied nor dissatisfied - fairly dissatisfied - very dissatisfied	10.0 8.1 5.9 4.5 2.4	0.1 0.1 0.1 0.1	10.0 8.1 5.7 4.2 2.3	0.1 0.1 0.1 0.1	10.0 8.1 5.8 4.3 2.3	0.0 0.0 0.1 0.1
O-SLu/g/sq/v/5/c	How satisfied are you with your life in general? - completely satisfied - satisfied - unsatisfied nor satisfied - unsatisfied - completely unsatisfied	10.0 8.7 6.2 4.3 1.8	0.1 0.1 0.1 0.1	10.0 8.5 5.9 4.0 1.5	0.1 0.1 0.1 0.1	10.0 8.6 6.0 4.2 1.6	0.0 0.1 0.1 0.1
O-HL/g/sq/v/7/a	If you were to consider your life in general, how happy or unhappy would you say you are, on the whole? - completely happy - very happy - fairly happy - neither happy nor unhappy - fairly unhappy - very unhappy - completely unhappy	10.0 9.2 7.6 5.8 4.6 3.0 1.1	0.1 0.1 0.0 0.1 0.1 0.1	10.0 9.2 7.5 5.6 4.3 2.8 1.1	0.1 0.1 0.1 0.1 0.1 0.1	10.0 9.2 7.5 5.7 4.4 2.9 1.1	0.0 0.0 0.0 0.0 0.0 0.0

The English translation is also given in this table, but has of course not been presented to the Dutch judges.