

# Chapter 2

## More Terminological and Methodological Problems in Measuring Happiness, Life Satisfaction and Well-Being: Some First Empirical Results

Anna Maffioletti, Agata Maida and Francesco Scacciati

### 2.1 Introduction

In spite of the abundance of literature on happiness economics (and surroundings, such as life satisfaction and well-being),<sup>1</sup> we believe that quite a number of terminological and methodological problems have not been solved yet: we will cope with some of them in this study.

### 2.2 Terminological Problems

In the literature, some terminological problems are present within languages and some others between languages.

- (a) As for those within languages one of the main questions is: are the notions of ‘happiness’, ‘life-satisfaction’ and ‘well-being’ substantially equivalent (as is taken for granted in most literature) or not? And, if not, are there some factors that systematically determine subjects’ different self-valuations when asked about their feeling happy or satisfied or well off? Moreover, are we sure that all researchers mean the same thing when using the same word? And that all respondents understand the same thing when facing the same word? And, if so, is it the same thing that researchers mean?
- (b) As for the problems between languages they regard translation. Most of the literature on happiness economics is, obviously, written in English, but most of the countries in which the questionnaires are distributed are not

---

<sup>1</sup> A very good survey and comments on literature coping with happiness, life satisfaction and well-being economics is in OECD Guidelines on Measuring Subjective Well-being (April 2013). Most of the methodological and terminological problems that are coped with in our chapter are discussed in this OECD issue: the comments show that for most of them no really conclusive argument has yet been reached. See also: Veenhoven (2010).

English-speaking countries. Therefore, there is a serious risk that something may be *lost in translation*. As for Italian, the keyword is itself problematic. All dictionaries will present, as first translation of the word *happiness*, the word ‘felicità’. But probably what is meant by *happiness* in English is not what most Italians mean by the word *felicità* (and the same may happen for Spanish, French, etc.). If asked, in an open question, to speak about the concept of ‘felicità’, quite a number of Italians would answer citing the old proverb ‘la felicità non è di questo mondo’, whereas probably very few English speaking subjects would answer ‘happiness does not belong to this world’. The fact is that the word *felicità* represents something extremely intense and absolute, that those that believe in god and in the immortality of the soul would say can be everlastingly enjoyed only in heaven and, in any case, most people would say that it can be enjoyed in life only in a few moments or, at most, for very short periods. On the contrary, happiness can be enjoyed even for ‘small’ events and for very long periods, also life-long, if you are lucky enough. Probably the word *happiness* may be better translated into Italian with the word ‘contentezza’ (and, therefore, *happy* with ‘contento’). There are no problems in translating *life satisfaction* in ‘soddisfazione per la propria vita’, whereas there is no way to translate the word *well-being*, unless you ascribe to it only its material and economical sense. In this case the translation would be ‘benessere’, but this is not what is meant in this field of research. Probably ‘qualità della vita’ (quality of life) is the best translation for *well-being*.

### 2.3 Methodological Problems

Some methodological problems are just as important and difficult to be solved as those discussed above. One of the main methodological problems deals with the possibility of assigning cardinal values to answers when questions are expressed in a numerical scale. In particular, often, psychologists accept the cardinal principle, whereas economists seem to prefer the more limiting but also more reliable ordinal comparability of answers. The question seems not to be amenable to a definitive solution. But even if the acceptability of the cardinal hypothesis could be proved and therefore would prevail, still quite a number of methodological problems would remain.

- (a) A first, a very important question that needs an answer is the following. ‘Do respondents, *facing a numerical scale*, first mentally answer the question *how happy are you?* (or something similar) in a verbal scale and then “translate” their answer into numbers?’ That is, do they first place their own valuation somewhere on a verbal scale, for instance going from ‘very unhappy’ to ‘very happy’, that has just as many steps as the numerical one and then give their numerical answer at the corresponding level or do they skip this step and respond directly in numerical terms? We thought this problem may be coped

with by asking the same subject the same question twice (at a good distance one from the other), the first time asking her/him to answer on a verbal scale and the second time on a numerical scale (half of the times the other way round).

If the answers are not significantly different, then the question at stake still stands unsolved (although the hypothesis of the ‘translation’ of concepts into numbers is strengthened), but if they are significantly different, this means that there is no ‘translation’, and therefore a second question pops up: which of the two answers is more reliable? We believe that most human beings would be more at ease, when asked ‘how happy are you?’, if they could choose their answer, say on a 7-step scale, between ‘very unhappy’, ‘unhappy’, ‘rather unhappy’, ‘neither unhappy nor happy’, ‘rather unhappy’, ‘happy’ and ‘very happy’, rather than between 1, 2, 3, 4, 5, 6 and 7, even if explained that 1 corresponds to the lowest level possible and 7 to the highest. Moving from this first methodological problem, we picked out two more of them.

- (b) Now comes the second question. ‘Is the wording adopted in a verbal scale relevant?’ More precisely, sticking to the ‘happiness example’, do respondents answer in the same way if they are asked to value their degree of happiness on a ‘*bipolar scale*’ going from ‘very unhappy’ to ‘very happy’ and on a ‘*unipolar scale*’ going from ‘not at all happy’ to ‘very happy’?

We believe that, if verbal scales are chosen, bipolar scales should be preferred. Let us take, for example, the international ranking presented in a survey by Adrian White in 2007. Denmark—the happiest country in the world, relying on his calculations—scores 2.73 whilst Burundi, in the last position, scores 100. (‘Scores have been presented from an index baseline of 100 for ease of comparison.’). Denmark therefore scores 273 times as much as Burundi: more or less on a 1–10 scale, 8.5 versus more or less 3. But three what? Three positive points: this means that unhappiness has been arbitrarily cancelled from the face of the Earth.

- (c) If you consider the above mentioned as a conclusive argument, does it lead necessarily to the convenience of adopting the bipolar structure also in numerical scales? That means, should we use scales going from negative to positive values (i. e. in an 11 point scale, from  $-5$  to  $+5$ )? Logically, the answer to this question should be ‘yes’, especially because of the ambiguity, in a unipolar numerical scale, of the first half of it, up to the middle point.

This aspect is present and discussed in detail in the above-mentioned OECD issue, as for verbal scales: *‘In a unipolar format, the scale midpoint is intended to represent a moderate amount of the variable of interest, whereas in a bipolar form at the mid point is intended to represent a more neutral territory in between the two opposing constructs: this should, in theory, have significant consequences for the meaning of the scale points. For example, a score of 0 on the unipolar scale above implies the absence of happiness, whereas a score of 5 implies a moderate amount of happiness. Conversely, on the bipolar scale, a score of 0 should denote complete unhappiness, a score of 5 implies the respondent is neither happy nor unhappy, and a score around 7 or 8 would imply a moderate amount of happiness.’*

But the same problems stand for the whole of the first half of the positive numerical steps. Answering the question “how happy are you?” on a scale going, say, from 0 to 10 what does a ‘3’ mean? Some subjects answering 3 will want to express being moderately happy (because 3 is anyhow a positive value), others being rather unhappy (because 3 is less than the middle point, that they consider representative of a ‘neither—nor’ valuation).

A counterargument in disfavour of bipolar numerical scales is that the positive side and the negative side may not be comparable with each other (they still allow for sums and averages, but not for ratios). Another one is that quite a number of respondents may not be at ease with negative numbers and therefore tend to ignore them. The OECD issue quotes quite a number of publications where it is shown that the average value of self-reported happiness (or life satisfaction or well-being) is significantly higher when measured by bipolar numerical scales because the negative numbers are less often selected than low positive ones in a unipolar numerical scale.

- (d) Another important question deals with scales. ‘Are numerical scales reliable?’ We consider a numerical scale reliable if the answer given on it fairly corresponds to the answer given on a verbal scale with the same number of steps. For example, let us imagine that subject A is the representative individual and she is asked ‘how happy are you?’ We don’t know exactly what goes on in her brain in the following instants, but we know that she will answer using words like ‘rather unhappy’ or ‘very happy’ and so on (or something very similar). Let us say that, facing an 11-point verbal scale—‘neither happy nor unhappy’ is at the sixth step—she answered, with not much difficulty (we may presume) ‘rather happy’ which is at the eighth step. Now let us imagine that, after a while, she is asked again the same question, but now she must select a number corresponding to her valuation on a scale going from 0 to 11. She *should* select ‘8’. If most of the subjects do so, and therefore spot out on the numerical scale the number that is in the same position as the expression selected on the verbal scale, we can state that numerical scales are reliable. Of course a ‘7’ or a ‘9’ would leave some uncertainty about the numerical scale reliability, but certainly something less than ‘7’ or more than ‘9’ would not. Even a ‘corresponding average’ would not warrant for reliability: we would also, obviously, need a very low variance. And, here comes a connected question, if numerical scales turn out to be reliable, are bipolar scales more reliable than unipolar scales or vice versa? That is, which type of numerical scale ‘fits’ better with the answers given on the verbal scales?
- (e) The fifth, and last (as for this study) question. ‘Can self-valuation be the only criterion on which public decision making for improving public happiness, life satisfaction and well-being should rely, or should public policies rely on (or also on) objective factors?’ In favour of the first option stands the simple argument that nobody can know better than oneself whether she/he is happy or not and what would help feeling better. But there are some reasons in support of the opposite option.

A factor that (at least to our opinion) should be adopted as an objective criterion, besides (not *instead*, of course) self-valuation is life expectation. The unemployed and the ill can make their voice audible in self-reported valuation surveys, but the dead, of course, cannot. They simply vanish from the sample. You can ask somebody how happy she/he would be if she/he knew that the next day he is going to be fired or she is going to break her leg. And if that really happens (and it does happen to plenty of people) you may still ask them how happy they feel after the mishap has actually occurred. You can also ask somebody how happy she/he would feel if she/he knew that she/he was going to die the next day (probably the valuation would be 0 in a 0–10 scale and minus whatever in a negative–positive scale). But if she/he actually does die (and it does happen to over 150,000 people every day in the world), then the negative valuation simply disappears from the survey—differently from the other cases. All mishaps affect negatively the overall average of self-reported happiness, except for the worst one of all. Therefore we believe that changes over time in subjective valuations are objectively underestimated in presence of significant increases in life expectation. For this and for quite a number of other reasons, the use of both subjective and objective criteria seems to us the most fruitful method of analysis. Nevertheless, a big question mark about the correct mix would still stand, difficult to be answered.

## 2.4 A Research on Happiness, Life Satisfaction and Well-Being in Piedmont

By means of a face-to-face 1,241 interview survey—beyond assessing the levels of happiness, life satisfaction and well-being perceived by the population of Piedmont and identifying the factors that determine them—we took the opportunity offered by a grant from the Regional Government of Piedmont to the Department of Economics of the University of Torino, to try to solve some of the problems discussed above.

Following the result of a pre-test, we decided to choose a bipolar verbal 7-step scale and both bipolar and unipolar 7-step numerical scales, because:

- (a) most of the literature adopts 11-point scales and we wanted to see if something different would show up with a different scale;
- (b) we noted that, beyond the 7 steps, the differences in a verbal scale between each step and the following were rather difficult to be appreciated: therefore, if we accept the ‘mental translation hypothesis’ (i.e. from concepts, verbally thought of, into numbers), also the difference between, say, a 2 and a 3 or between a 7 and an 8, should be equally difficult to be appreciated on an 11-step numerical scale;
- (c) we believed that, if valuations are ‘translated’ from words into numbers, it may be more difficult to spot out a ‘4’ in a unipolar scale corresponding to a

‘neither unhappy nor happy’ valuation rather than a ‘zero’ on a bipolar scale, in spite of the fact that, in the literature, practically only positive numerical scales are adopted.<sup>2</sup> Therefore we adopted bipolar numerical scales;

- (d) we also adopted unipolar numerical scales to compare their results with the bipolar ones.

## 2.5 Some First Results

We did not even try to give an answer to all the problems we mentioned above, but we did try to answer some of them.

We coped with the terminological question if the terms ‘happiness’, ‘life satisfaction’ and ‘well-being’ are understood as synonymous—and with the linked question if they are determined by the same factors—and with the numerical scale question.<sup>3</sup>

- (a) To be able to face the first question, we simply asked each of the 1,241 interviewed subjects how they valued, for themselves, the level of each of these notions. There were two questionnaires, A and B (50 % of the sample each), identical except for the position of the questions on happiness, satisfaction and well-being that were placed in opposite order at the same distance one from the other. Each questionnaire contained 63 ‘main’ questions and 69 more ‘sub-questions’, and therefore it was easy to distance enough the questions on these three notions so as to exclude, or at least limit, the influence that the first answer may have on the following ones.

On the verbal scale, we find significant mean differences between happiness (4.97) and well-being (4.83) and between well-being and life satisfaction (4.92), whereas the notions of happiness and satisfaction do not show significant differences. But, even more interestingly, the determinants of the three notions appear to be different, and, within this phenomenon, there are also significant differences between genders.

Income<sup>4</sup> turns out to be very significantly relevant for all the three notions, but more as for well-being than for life satisfaction and for happiness: moreover, in all three cases, as income grows, its marginal significance decreases. This confirms the results of Clark et al. (2008) that provide a mathematical explanation

---

<sup>2</sup> A numerical scale with only negative values (from  $-7$  to  $-1$ ) have been presented by Davern and Cummins (2006) to measure dissatisfaction, (which is not really the same thing as our bipolar  $-3$  to  $+3$  scale to measure dissatisfaction and satisfaction on the same scale).

<sup>3</sup> First elaborations of the collected data are in Maffioletti et al. (2013)

<sup>4</sup> In spite of assured anonymity, almost half of the samples did not answer the question on family income. As a consequence we performed two sets of analysis: in the first, only individuals who provided information on family income are considered, in the second the whole sample is considered, not including the income variable among regressors, but adding, as a row proxy of income, self-reported satisfaction of subject’s family on its overall economic condition.

of the Easterlin paradox<sup>5</sup> (i.e. the lack of correlation between income and happiness) in time-series analyses, and those of many recent studies—both for countries all together<sup>6</sup> and for Italy specifically<sup>7</sup>—that show that happiness is significantly positively related to income in cross-section analyses.

We also find that the factors that determine the positive side of the scale are different from those that determine the negative side: surely employment/unemployment and good/bad health seem to have this type of effect. Therefore—if the final results of our research are supposed to be also prescriptive for public policies<sup>8</sup>—it is necessary to understand whether these policies, to be as much effective as possible, should act (as far as possible) more on the factors that determine the positive side of self-reported happiness/life satisfaction/well-being—that seem to be more difficult to be affected—or on those that determine the negative side (Regional policies can affect strongly both health and employment).

Moreover, it is not clear whether factors valued positively and factors valued negatively, when adopting a self-valuation technique, counterbalance each other, or whether a negative valuation about a very important factor prevails on the valuation of any other in determining the ‘all together’ valuation, independently of the personal weight ascribed<sup>9</sup> to the factor(s) valued negatively. And, the other way round, does a negative valuation of one’s life all together determine negative valuations of single factors? We have a hint of this coming from a parallel research led from our colleagues in Alessandria (based on the same questionnaires): high valuations of happiness and well-being are associated with positive valuations of public services: in which direction goes the cause–effect relation?

(b) As far as scales are concerned, the first question that shows up is if bipolar numerical scales are more reliable than unipolar numerical scales. It turns out that:

(b1) the mean answer as for *life satisfaction* expressed in the verbal scale was 4.92, whereas the mean answer in the numerical bipolar scale was 5.08.

(b2) the mean answer to the question on *happiness* expressed in the verbal scale was 4.97 while it was 4.83 when the numerical unipolar scale is used.

*T* test shows that both mean differences are statistically significant.

On the contrary, no significant difference emerged either between self-reported *well-being* in a verbal scale (mean 4.83) and in an numerical unipolar scale (mean 4.84) in questionnaire A, or between self-reported well-being in a verbal scale (mean 4.86) and in an numerical bipolar scale (mean 4.75) in questionnaire B.

These results are rather mixed and there is no conclusive argument neither for the reliability of numerical scales, nor for the preference that should be conferred to one type of scale or to the other: in all cases differences are (above

<sup>5</sup> See Easterlin (1973, 1974, 2005).

<sup>6</sup> Up to this moment, Stevenson and Wolfers (2013) is one of the most recent and convincing.

<sup>7</sup> See Scoppa and Ponzio (2008), ISTAT (2013).

<sup>8</sup> See Layard (2005, 2006).

<sup>9</sup> Self-reported, when answering questions such as “How important is it for you...”.

or below) at the borderline of significance. But it is interesting to note that, when the differences are significant, the direction is the expected one: the unipolar scale underestimates valuations because of the mixed understanding of the low steps, up to the middle one (subject A may believe that even a '3' could represent correctly her being 'rather happy', being, anyhow, a positive value); the bipolar scale overestimates valuations, because of the fact that quite a number of respondents may not be at ease with negative numbers and therefore tend to ignore them and therefore consider only the numbers from 0 upwards as possible answers.

## 2.6 Conclusions

On the basis of our results we may conclude that:

- (1) People do not consider 'happiness', 'life satisfaction' and 'well-being' as synonymous, and this is proven not only because the mean of self-valuations is significantly different, but also because the determinants of these three notions are different.
- (2) The factors that determine the positive side of the scale are different from those that determine the negative side: some factors, such as employment and health determine self-reported happiness, life satisfaction and well-being only if negatively reported (i.e. unemployment and bad health), but not if they are reported positively. Public policies are probably bound to be more effective if intervening on the factors that determine unhappiness rather than on the factors that determine happiness, because the former ones seem to be more suitable to be affected by public policies.
- (3) Income has a significant effect on self-reported happiness, life satisfaction and well-being, as spotted out by the prevailing literature in cross-section analysis, is confirmed. The marginal decrease of its significance is also confirmed.
- (4) The use of different scales is relevant, but we did not find a conclusive answer neither to the question if numerical scales are fully reliable, nor—in case of a positive answer to this question—which, between a unipolar and a bipolar scale, is more reliable than the other.

## References

- Clark, A., Frijters, P., & Shields, M. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, 46(1), 95–144.
- Davern, M., & Cummins, R. A. (2006). Is life dissatisfaction the opposite of life satisfaction? *Australian Journal of Psychology*, 58(1), 1–7.
- Easterlin, R. (1973). Does money buy happiness? *The Public Interest*, 30, 3–10.



- Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz* (pp. 89–125). New York and London: Academic Press.
- Easterlin, R. (2005). A puzzle for adaptive theory. *Journal of Economic Behavior and Organization*, 56, 513–521.
- ISTAT. (2013). *Rapporto annuale 2013 §4*.
- Layard, R. (2005). *Happiness—lessons from a new science*. London: Penguin.
- Layard, R. (2006). Happiness and public policy: A challenge to the profession. *The Economic Journal*, 116, 24–33.
- Maffioletti, A., Maida, A., & Scacciati, F. (2013). *Survey design and response analysis: A study on happiness, life satisfaction, and well-being in piedmont, a region of Italy*. LABORatorio Riccardo Revelli, Working Papers Lrr 131.
- OECD. (2013). *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/9789264191655-en>; <http://worlddatabaseofhappiness.eur.nl/hap>
- Scoppa, V., & Ponzo, M. (2008). An empirical study of happiness in Italy. *The B.E. Journal of Economic Analysis and Policy*, 8, Article 15.
- Stevenson, B., & Wolfers, J. (2013). *Subjective well-being and income: Is there any evidence of satiation?* NBER Working Paper No. 18992, April.
- Veenhoven, R. (2010). Life is getting better: Societal evolution and fit with human nature. *Social Indicators Research*, 97, 105–122.