Well-being Changes from Year to Year: A Comparison of Current, Remembered and Predicted Life Satisfaction

Marcus Klemm

Accepted: 13 October 2021 / Published online: 8 November 2021
© The Author(s) 2021

Abstract
I study yearly changes in personal well-being combining data on current, retrospective and prospective life satisfaction from the German Socio-Economic Panel. Predicted and remembered changes in life satisfaction are both positive on average and match well, whereas the average year to year-change inferred from reports of current life satisfaction is negative. Retrospective assessments of past well-being are strongly influenced by current life satisfaction, significantly related to past life satisfaction and linked to past predictions of current satisfaction. Due to different problems related to the ordinal measurement scale, changes in subjective reference systems and recall ability, the analysis overall suggests that direct reports of intertemporal changes provide valuable additional information for the analysis of individual well-being.

Keywords Life satisfaction · Intertemporal comparison · Response shift · Memory bias · Forecast error · Adaptation

JEL Classification A12 · C18 · I31

1 Introduction

Over the last decades, economists have become increasingly interested in the study of individual life satisfaction, subjective well-being or happiness (see, e.g., the reviews by Frey and Stutzer 2002; Di Tella and MacCulloch 2006; Dolan et al. 2008; Clark 2018). These survey data can deliver new insights and offer ways to test economic theories and to advise or evaluate public policy (Stutzer and Frey 2010; Di Tella and MacCulloch 2006), but there are also reasons to be skeptical about their informational content (e.g., Schwarz

1 As common in the economic literature, I use the terms life satisfaction, well-being and happiness interchangeably.

I thank two anonymous referees, David Bartram and Martin Kukuk for very helpful comments and discussions.

Marcus Klemm
marcus.klemm@fhws.de

1 University of Applied Sciences Würzburg-Schweinfurt, Münzstr. 12, 97070 Würzburg, Germany
and Strack 1999; Bertrand and Mullainathan 2001; Bond and Lang 2019; Schröder and Yitzhaki 2017). My study contributes to the research on life satisfaction by combining longitudinal data on current, retrospective and prospective assessments of personal well-being from the first four waves of the German Socio-Economic Panel (GSOEP). Thereby, yearly changes in life satisfaction can be looked at from three different perspectives: (1) by comparing current assessments from subsequent years, (2) ex-ante by comparing current and predicted satisfaction, (3) ex-post by comparing current and remembered satisfaction. The study aims to provide first descriptive evidence on systematic links between these three global assessments of individual life satisfaction.

Since life satisfaction data are subjective assessments of the personal situation, comparison standards, i.e., the reference system, for such assessments play an important role for their analysis and the interpretation of empirical findings. For instance, the context in which life satisfaction questions are asked (Pavot and Diener 1993; Oishi et al. 2003; Deaton and Stone 2016) or the relative position in one’s social and economic environment (e.g., Kapteyn et al. 1978; Clark et al. 2008b; Senik 2009; D’Ambrosio and Frick 2012) can affect reports of life satisfaction. When studying the development of well-being over time based on recurring assessments of the personal situation at different time points, these assessments are subject to changes in the reference system as discussed in the literature on response shift (Schwartz and Sprangers 1999b) and preference drift (van Praag 1971; Groot and Maassen van den Brink 2000) which can lead to (hedonic) adaptation or biases in such assessments (Groot 2000). There is ample evidence from concurrent well-being data concerning adaptation and habituation of individuals to changing life circumstances over time, in particular with regard to income (e.g., Kapteyn et al. 1978; Clark et al. 2008b; Di Tella et al. 2010; Clark 2018), but also regarding poverty (Clark et al. 2016) marriage, divorce, children or unemployment (Clark et al. 2008a; Clark and Georgellis 2013).

Retrospective and prospective life satisfaction data offer additional perspectives on intertemporal comparisons. The reference system for the simultaneous assessment of current, retrospective and prospective life satisfaction most likely differs from the reference systems for the current well-being-assessments at different time points. A divergence of findings might be interpreted as “mispredicting and misremembering” (Smith et al. 2008), but it could also reflect different measurement concepts. Schwandt (2016) compares expectations about life satisfaction in five years to the corresponding life satisfaction reports five years later and concludes that younger people are overly optimistic whereas older people are overly pessimistic. Prati and Senik (2020) compare remembered satisfaction to current assessments and conclude that people ex-post overstate changes in well-being compared to the development implied by their corresponding current reports. Kaiser (2020) uses the same data from the British Household Panel Study (BHPS) as Prati and Senik (2020) and finds the same descriptive patterns which can be due to both scale shifts or memory biases. Hagerty (2003) reports that the majority of individuals assesses their current situations to be better than in the past based on a meta-analysis of 71 direct intertemporal assessments. My findings match these results on intertemporal changes in life satisfaction in that

---

2 The concept of response shift can be traced back to research on educational interventions and organizational change in the 1970’s (Schwartz and Sprangers 1999b, a). The terms response shift or preference drift can subsume different phenomena such as recalibration of personal well-being scales, redefinition of the importance of different well-being determinants or entire changes of well-being concepts (see, e.g., for further discussion Schwartz and Sprangers 1999b; Ubel et al. 2010; Sprangers and Schwartz 2010; Odermatt and Stutzer 2019.)
recalled and predicted changes in well-being are more positive than implied by comparing the corresponding current assessments. Furthermore, the simultaneous look at current, retrospective and prospective assessments suggests that predictions and subsequent recall of life satisfaction are systematically linked and that these data carry additional informational value for the analysis of well-being developments over time. The next section presents the data and the descriptive empirical analysis in detail. The third section provides further discussion and concludes.

2 Data and Empirical Analysis

The first four waves of the German Socio-Economic Panel (GSOEP) from 1984 to 1987 offer individual-level data of more than 10,000 respondents on annual changes in personal life satisfaction. The GSOEP was started as a longitudinal survey in 1984 (for a detailed description, see Wagner et al. 2007). Since then, a typical question on life satisfaction has been surveyed each year: “And finally, we would like to ask you about your satisfaction with your life in general. Please answer by using the following scale, in which 0 means totally unhappy, and 10 means totally happy. How happy are you at present with your life as a whole?” In the following, the answers to this question are labeled “current satisfaction” $CS_{it}$, i.e., life satisfaction of individual $i$ in year $t$ as reported in year $t$. Until 1987, the question that directly followed was: “How happy were you a year ago with your life?” The answers to this question are labeled “retrospective satisfaction” $RS_{it}$, i.e., life satisfaction in year $t - 1$ as reported in year $t$. Finally, the survey participants were asked: “And what do you think it will be like in a year’s time?” These answers are labeled “prospective satisfaction” $PS_{it}$, i.e., life satisfaction in year $t + 1$ as expected by individual $i$ in year $t$.

These variables permit the calculation of changes in life satisfaction between years $t - 1$ and $t$ in three ways:

\[
\Delta CS_{it} = CS_{it} - CS_{it-1}, \tag{1}
\]

\[
\Delta RS_{it} = CS_{it} - RS_{it}, \tag{2}
\]

\[
\Delta PS_{it} = PS_{it-1} - CS_{it-1}. \tag{3}
\]

Depending on the question asked, respondents can be expected to focus more on temporal, personal or social comparisons (see Dubé et al. 1998; Schwarz and Strack 1999; McBride 2010, for more detailed discussions). The latter two measures probably put relatively more weight on intrapersonal intertemporal comparisons because the time dimension is explicit in the questions. Since these assessments are made at the same time, social reference groups and aspirations should be identical for $CS_{it}$ and $RS_{it}$ as well as for $CS_{it-1}$ and $PS_{it-1}$, or at least much more similar than for the current assessments $CS_{it}$ and $CS_{it-1}$. Differences between $\Delta CS$, $\Delta RS$ and $\Delta PS$ can be due to different comparison standards, but also due to recall problems, memory bias or forecast errors, e.g., people being too optimistic on average (Odermatt and Stutzer 2019; Schwandt 2016; Frijters et al. 2009). The accuracy of forecasts and recall naturally declines the longer the time horizon. Gibson and Kim (2010)

---

3 Dubé et al. (1998), Steffel and Oppenheimer (2009) and Senik (2009) find that intrapersonal are more important than interpersonal comparisons, which should limit the difference between the three measures.
note that especially long-term retrospective recall poses a serious problem to data accuracy, and Hagerty (2003) suggests that recall only becomes difficult after 5 years. Here, the time period between interviews is approximately 1 year matching the time horizon of the survey question quite well (see Table 1). For more than 90% of the interviews, the time between surveys is between 9 and 15 months.4

Table 1 shows the average current, prospective and retrospective satisfaction scores along with the corresponding average yearly changes. In order to create a consistent sample for which the yearly changes in satisfaction refer to the same time periods for all measures, only individuals for which all three satisfaction measures are available in at least two consecutive surveys are included in the analysis.5 On average, current satisfaction has declined by 0.15 points annually between 1984 and 1987. In contrast, prospective and retrospective reports show average annual increases by 0.15 and 0.07 points, respectively. Such differences between current, prospective and retrospective, and also between personal and more global/societal assessments are common in the literature (Hagerty 2003; Gandelman and Hernandez-Murillo 2009; Deaton and Stone 2013; Schwandt 2016; Prati and Senik 2020). The correlation of $RS_{t}$ and $CS_{t-1}$ amounts to 0.5, the correlation of $PS_{t-1}$ and $CS_{t}$ to 0.4, which suggests that the intertemporal assessments capture a significant part of past and future current satisfaction (Krueger and Schkade 2008; Pavot and Diener 1993), but t-tests strongly reject the hypothesis that $\Delta CS$, $\Delta RS$ and $\Delta PS$ are equal on average.

### Table 1 Descriptive statistics

|                                | Mean  | Std. dev. | Min  | Max  | Observations |
|--------------------------------|-------|-----------|------|------|--------------|
| Current satisfaction in t      | 7.21  | 1.98      | 0.00 | 10.00| 29,894       |
| Current satisfaction in t-1    | 7.36  | 2.02      | 0.00 | 10.00| 29,894       |
| Prospective satisfaction from t-1 about t | 7.51  | 1.98      | 0.00 | 10.00| 29,894       |
| Retrospective satisfaction from t about t-1 | 7.14  | 1.97      | 0.00 | 10.00| 29,894       |
| $\Delta$ Current satisfaction | −0.15 | 2.07      | −10.00 | 10.00| 29,894       |
| $\Delta$ Prospective satisfaction | 0.15  | 1.22      | −10.00 | 10.00| 29,894       |
| $\Delta$ Retrospective satisfaction | 0.07  | 1.32      | −10.00 | 10.00| 29,894       |
| Months between interviews      | 11.20 | 1.76      | 5.00 | 21.00| 29,894       |
| Age                            | 42.89 | 16.64     | 16.00| 98.00| 29,894       |
| Female                         | 0.51  | 0.50      | 0.00 | 1.00 | 29,894       |

Note: The sample only includes individuals observed in at least two consecutive years. Therefore, the time index $t$ refers to the years 1985 to 1987, and the $\Delta$-variables which measure annual changes refer to the time period from 1984 to 1987. Source: GSOEP 1984–87, own calculations

4 Manzoni et al. (2011) provide a thorough discussion of prospective and retrospective survey design, including issues related to memory biases and recall period.

5 Therefore, retrospective satisfaction reported in 1984 about 1983 does not enter the analysis because neither current satisfaction in the year 1983 nor prospective satisfaction reported in 1983 about 1984 are available. Likewise, prospective satisfaction reported in 1987 about 1988 does not enter the analysis because retrospective satisfaction has not been asked in the GSOEP from 1988 onwards. Current satisfaction in year $t$ comes from the years 1985 to 1987 while current satisfaction in year $t$-1 refers to the years 1984 to 1986. Including $RS_{1984}$, $PS_{1987}$, $CS_{1984}$ among current satisfaction in year $t$ and $CS_{1987}$ among current satisfaction in year $t$-1 does not substantially change the reported statistics: Including all available data, the mean values are $CS = 7.28$, $PS = 7.45$ and $RS = 7.17$. 

 Springer
While the three measures of yearly changes in life satisfaction are clearly not identical, $\Delta RS$ and $\Delta PS$ appear to match much better: $\Delta RS$ and $\Delta PS$ are identical in 52% of all observations, $\Delta CS$ and $\Delta PS$ in 27% and $\Delta CS$ and $\Delta RS$ in 18%. Allowing for a 1 point-difference, the congruence of the three measures amounts to 77%, 60% and 39%, respectively. A significant part of the better match between $\Delta RS$ and $\Delta PS$ can be traced back to a large fraction of recalled or predicted life satisfaction changes of 0: 67% and 64%, respectively. In contrast, only 29% of satisfaction changes measured by $\Delta CS$ are 0. The significantly larger variation of $\Delta CS$ is also mirrored by the standard deviations of the three measures in Table 1.

The comparison of retrospective and current data provides a similar picture as in Prati and Senik (2020) and Kaiser (2020) with regard to people over-stating their past life satisfaction development. The left panel of Fig. 1 shows the predicted and remembered changes in life satisfaction by the observed change in current satisfaction. In qualitative terms, all three measures coincide, but especially large negative developments are typically not expected. For all measures, reporting problems arise due to the bounded scale. In particular, individuals at the lower and upper ends of the scale cannot predict a further decline or increase in life satisfaction, and such changes cannot be identified from the current assessments. For example at the upper end of the scale, average $\Delta CS$ and $\Delta PS$ are null or negative by definition because values higher than 10 cannot be reported. The right panel of Fig. 1 shows that the retrospective assessments suggest positive well-being developments among people who report the same higher levels of life satisfaction in two consecutive years. Among those who report a current satisfaction of 10 in two consecutive years, $\Delta RS$ is 0.2 on average ($n=1,743$). At the lower end of the scale, $\Delta RS$ is -0.8 on average for those with current satisfaction below 5 and $\Delta CS = 0$ ($n=200$). For all those with $\Delta CS = 0$, $\Delta PS$ decreases along the life satisfaction scale while $\Delta RS$ increases. This also points to another issue with the discrete ordinal measurement of unobservable latent continuous satisfaction: In order to change one’s current assessments, stronger impulses might be necessary as assessments move closer to the more extreme values of the scale (see also Kaiser 2020).6

Figure 2 plots the yearly change in life satisfaction by gender and age groups. Overall, the annual changes in life satisfaction are slightly larger in absolute terms for females. While the negative average change in current life satisfaction $\Delta CS$ is not statistically distinguishable between females and males, females predict and remember significantly higher

---

6 Since average life satisfaction is higher than 7 and more than 50% of all reports between 6 and 8 (approx. 80% between 5 and 9), values below 5 can be regarded as very low; values of 10 and 9 as very high.
satisfaction differences $\Delta PS$ and $\Delta RS$. Consequently, the difference between the three satisfaction measures turns out larger for females. Similar to Schwandt (2016), younger people are more optimistic and older people more pessimistic about their future well-being. A similar pattern is visible for remembered changes in well-being: People report well-being improvements when young and well-being reductions when old. From midlife onward, predicted and especially remembered changes are close to zero on average. For the two oldest age groups, all three measure point into the same negative direction. One explanation for this pattern could be stronger changes in reference systems, i.e., individuals rescale their life satisfaction assessments, until roughly the late 30’s. Hence, it is possible that individual well-being increases on average at young age, while reported current satisfaction scores point into the opposite direction. Since younger individuals report higher levels of current satisfaction (more than half of those under the age of 30 report a current life satisfaction of 8 or higher), the issues related to the discrete ordinal measurement discussed above might also play a role.

The discussed patterns could be due to people generally being too optimistic when young and overstating their own past development at the same time (memory bias), e.g., for self-appraisal. In order to investigate the potential systematic links between the different life satisfaction assessments further, I start from Equations (2) and (3), assuming that $\Delta RS$ and $\Delta PS$ should be equal in the absence of systematic or random errors, e.g., memory or prediction biases or changes in the subjective reference system. Rearranging yields the following regression equation for retrospective life satisfaction in year $t-1$ as reported by an individual $i$ in year $t$:

$$RS_{it} = \alpha CS_{it-1} + \beta CS_{it} + \gamma PS_{it-1} + \eta_i + \epsilon_{it}.$$  

This equation includes the main global determinants of retrospective life satisfaction: past current satisfaction $CS_{it-1}$ about a year ago, the current level of satisfaction $CS_{it}$ which could be regarded as the anchor for the intertemporal assessments and the past prediction of current satisfaction $PS_{it-1}$. In addition to a random error term $\epsilon_{it}$, I include an individual fixed effect $\eta_i$ because certain personality traits (e.g., optimism or extroversion) have the potential to affect retrospective assessments similar to other life satisfaction assessments

---

7 The observed age pattern is quite similar for females and males.

8 From $\Delta RS_{it} = \Delta PS_{it}$ follows $CS_{it} - RS_{it} = PS_{it-1} - CS_{it-1}$ and $RS_{it} = CS_{it-1} + CS_{it} - PS_{it-1}$. In the absence of any random or systematic differences, we would expect $CS_{it-1} = RS_{it}$, $CS_{it} = PS_{it-1}$ and $\Delta CS = \Delta PS = \Delta RS$. As Table 1 and, e.g., Schwandt (2016) or Kaiser (2020) show, this does not appear to be the case.
(Ferrer-i-Carbonell and Frijters 2004). In case of an intertemporally stable reference system and the absence of any other systematic factors (e.g., recall problems), $\alpha$ should be 1, while $\beta$ and $\gamma$ should be 0, i.e., $RS_{it}$ should only depend on $CS_{it-1}$, the well-being level to be recalled.

While this simple regression analysis is not deeply based on psychological and economic theory, it provides a general description how the global assessment of past, current and future well-being are linked. The panel nature of the data allows me to control for unobserved time-invariant individual factors summarized by $\eta_i$. Furthermore, the regressions include year dummies and the time difference in months between the two surveys from which the satisfaction assessments stem. Table 2 reports the results of linear random and fixed effects regressions, a fixed effect ordered logistic regression by Baetschmann et al. (2015, 2020) and a Gini regression (Schaffer 2015), standard errors adjusted for clustering on individuals in parentheses (except Gini regression). All regressions include year dummies. Significance levels: *10% **5% ***1% Source: GSOEP 1984–87, own calculations

Table 2 Determinants of retrospective life satisfaction reported in $t$ about $t−1$

|                      | Linear random effects | Linear fixed effects | Fixed effects ordered logit | Gini regression |
|----------------------|-----------------------|----------------------|-----------------------------|-----------------|
| Current satisfaction in $t−1$ | 0.209***               | 0.170***             | 0.297***                    | 0.201***        |
|                       | (0.010)               | (0.014)              | (0.023)                     | (0.007)         |
| Current satisfaction in $t$  | 0.700***               | 0.640***             | 0.847***                    | 0.734***        |
|                       | (0.007)               | (0.010)              | (0.026)                     | (0.004)         |
| Prospective satisfaction from $t−1$ | -0.078***             | -0.119***            | -0.179***                   | -0.076***       |
|                       | (0.009)               | (0.013)              | (0.022)                     | (0.007)         |
| Months between interviews | -0.012**               | -0.016***            | -0.038***                   | -0.010*         |
|                       | (0.005)               | (0.006)              | (0.012)                     | (0.005)         |
| Overall/Pseudo R2/GR     | 0.623                  | 0.619                | 0.355                       | 0.822           |
| Observations            | 29,894                 | 29,894               | 24,435                      | 29,894          |
| Individuals             | 11,559                 | 11,559               | 8438                        | 11,559          |

Table reports coefficients from linear random and fixed effects regressions, a fixed effects ordered logit model (Baetschmann et al., 2020) and a Gini regression (Schaffer, 2015), standard errors adjusted for clustering on individuals in parentheses (except Gini regression). All regressions include year dummies. Significance levels: *10% **5% ***1% Source: GSOEP 1984–87, own calculations

9 The cardinal treatment of the ordinal satisfaction measures is not without problems (Bond and Lang 2019; Schröder and Yitzhaki 2017). In my case, the dependent and explanatory variables are essentially measured on the same scale and monotonic transformations would affect the variables in similar ways. Furthermore, the so-called LMA curves do not intersect the horizontal axis and the signs of the regression coefficients are unlikely to be affected by the cardinal treatment of the satisfaction variable (Schröder and Yitzhaki 2017). Furthermore, Kaiser and Vendrik (2020) argue that sign reversals are not likely to be a serious concern in applied empirical research on life satisfaction, but that relationships between variables might be heterogeneous across the satisfaction scales.

10 The results of the random effects regression are also robust to the inclusion of a female-dummy, age and age squared as additional variables. Excluding the year dummies and the months between interviews does not affect any estimation results, either.
determinants of current, prospective and retrospective life satisfaction is beyond the scope of this analysis.

Retrospective satisfaction is significantly related to all of the three other satisfaction measures. The positive coefficient of life satisfaction in the reference period $CS_{t-1}$ creates confidence that retrospective assessments carry significant information about the survey participants’ past well-being. But as $\alpha$ is far from 1, recall error is very likely, not least due to the imprecise time frame of “a year ago”. When limiting the sample to only those surveys that were conducted exactly 12 months apart, the coefficient increases slightly (first column of Table 3). The strongest correlation emerges between retrospective and current life satisfaction in the reporting period $CS_t$. This is unsurprising because the question about current satisfaction is asked directly before the question about retrospective satisfaction. Thus, the former assessment probably serves as the anchor for the latter. This also suggests that the retrospective assessment is less about the recall of the satisfaction score reported a year ago, but more about an assessment of life a year ago relative to the current situation.

In all regressions, a significant negative correlation between the past prediction of current satisfaction and the current recall of past satisfaction is found. The size of the coefficient from the linear regressions ranges from -0.04 for the sample excluding high and low satisfaction scores to -0.2 for the sample of individuals surveyed 12 months apart. The coefficients turn out similar across all age groups. This implies that positive predictions are related to future retrospective downgrades of life satisfaction assessments, and vice versa. If predicted and retrospective assessments were both relative to the corresponding current satisfaction scores, the assessments of $CS_{t-1}$ and $PS_{t-1}$ would be based on the

Table 3  Determinants of retrospective life satisfaction reported in t about t-1 based on different samples

| Samples                  | 12 months | Balanced panel | Excluding panel extremes | Age 30 or younger | Between 30 and 50 | Age 50 or older |
|--------------------------|-----------|----------------|--------------------------|-------------------|-------------------|----------------|
| CS in t-1                | 0.244***  | 0.167***       | 0.058***                 | 0.203***          | 0.175***          | 0.118***       |
|                          | (0.032)   | (0.014)        | (0.018)                  | (0.026)           | (0.022)           | (0.023)        |
| CS in t                  | 0.626***  | 0.646***       | 0.773***                 | 0.548***          | 0.626***          | 0.723***       |
|                          | (0.021)   | (0.010)        | (0.013)                  | (0.020)           | (0.018)           | (0.015)        |
| PS from t-1              | -0.205*** | -0.111***      | -0.036**                 | -0.138***         | -0.108***         | -0.097***      |
|                          | (0.030)   | (0.013)        | (0.017)                  | (0.025)           | (0.021)           | (0.021)        |
| Months between interviews| -0.009*   | -0.004         | -0.009                   | -0.032***         | 0.005             |
|                         | (0.005)   | (0.006)        | (0.012)                  | (0.009)           | (0.009)           |
| Within R2                | 0.372     | 0.395          | 0.521                    | 0.276             | 0.370             | 0.522          |
| Overall R2               | 0.618     | 0.627          | 0.642                    | 0.486             | 0.619             | 0.731          |
| Observations             | 11,634    | 25,674         | 10,528                   | 8,377             | 11,640            | 9,877          |
| Individuals              | 7,636     | 8,588          | 6,131                    | 3,711             | 4,688             | 3,945          |

Table reports coefficients from linear fixed effects regressions, standard errors adjusted for clustering on individuals in parentheses. Extreme values are defined as all current, remembered and predicted satisfaction levels lower than 5 or higher than 8. All regressions include year dummies. Significance levels: *10% **5% ***1% Source: GSOEP 1984–87, own calculations

11 The exclusion of high and low scores, and thus also larger changes in life satisfaction, greatly reduces the variation in the data and disregards cases that might be of particular interest for the study of intertemporal satisfaction assessments.
reference system of year \( t - 1 \) and the assessments of \( CS_t \) and \( RS_t \) would be based on the potentially different reference system of year \( t \). Since the regressions control for current life satisfaction in both years and individual fixed effects, the results provide evidence that the differences between the direct intertemporal assessments \( \Delta PS \) and \( \Delta RS \) compared to \( \Delta CS \) are most likely not simply due to misremembering and misprediction, but reflect additional information about well-being, its development and changes in the reference system. In particular, small improvements in personal well-being can more easily be reported in predictions and in retrospect, but such changes might not be large enough to induce change in the current assessments, especially towards the end of the life satisfaction scale (right panel of Fig. 1). This is compatible with hedonic adaptation, not necessarily with regard to larger shocks or life events (Clark et al. 2008a), but rather on a small scale and as a continuous process.

3 Discussion and Conclusion

Most empirical research on personal well-being is based on current subjective reports. My study adds to the literature by linking current, retrospective and prospective reports of personal life satisfaction. In contrast to satisfaction changes measured by comparing current assessments from two years, people overpredict future well-being ex-ante (Odermatt and Stutzer 2019; Schwandt 2016; Frijters et al. 2009) and overstate their well-being development ex-post (Prati and Senik 2020; Kaiser 2020; Smith et al. 2008). In contrast, predicted and remembered changes in life satisfaction match quite well. This pattern is particularly pronounced at young age, i.e., young people predict and recall improvements in their well-being whereas their current life satisfaction scores decline on average. This observation could be due to relatively strong changes in the reference system for the assessment of personal life satisfaction.

Odermatt and Stutzer (2019) point out for prediction errors that there are several possible methodological, psychological or behavioral explanations why different life satisfaction assessments can produce different empirical results concerning the development of individual well-being. Based on the same GSOEP data on current and remembered satisfaction as my analysis, they conclude that rescaling cannot fully explain the observed life satisfaction adjustments following particular life events like marriage or becoming unemployed (Odermatt and Stutzer 2019, Online Appendix 7). Additionally accounting for life satisfaction predictions, my analysis suggests that intertemporal assessments provide an informational value beyond that of current assessments alone. In particular, the former can help to address adjustments of the measurement scale and thereby shed more light on the influence of reference systems and thus identify better true changes in well-being and adaptation processes (see, e.g., Kaiser 2020; Odermatt and Stutzer 2019).

The discrete ordinal measurement of latent well-being and especially the cardinal treatment of such data must be seen critical (Kaiser and Vendrik 2020; Bond and Lang 2019; 12 Since intrapersonal and intertemporal comparisons must be considered important reference standards (Dubé et al. 1998; Steffel and Oppenheimer 2009; Senik 2009; D’Ambrosio and Frick 2012), general misremembering and misprediction would largely reduce the informational value of any well-being data, especially for the analysis of developments over time. If individuals are seen as the best judges of their own well-being (Frey and Stutzer 2002), then they should probably also be good judges of how their well-being changes.

12
Schröder and Yitzhaki (2017). Small changes in well-being might not be enough to induce changes in current assessments. The stimulus necessary for changing one’s assessment might vary across the life satisfaction scale. And positive like negative developments cannot be reported at the ends of the measurement scale. The comparison of current and intertemporal assessments provides evidence that these issues exist: Self-reported changes in well-being vary over the life satisfaction scale with the highest absolute values in the lower half of the scale. The remembered (predicted) change in life satisfaction is higher (lower) for high levels of satisfaction and lower (higher) for low levels of satisfaction. Prati and Senik (2020) attribute the latter finding to behavioral factors, e.g., strategic self-enhancement. I suspect that both aspects are relevant, but disentangling measurement problems and behavioral factors further is beyond the scope of this study and left to further research (see also Kaiser 2020).

While recall difficulties and memory biases probably limit the informational content of retrospective assessments of past well-being, current assessments might over-accurate short-lived life circumstances, mood and affect (Deaton 2012; Wunder 2012; Schwarz and Strack 1999). The GSOEP data show a strong correlation between the contemporaneous assessments of current and past life satisfaction. Since the retrospective and past assessments are in fact significantly correlated (although the time horizon is not perfectly defined) the former might be well suited to study well-being more globally, net of rather noisy short-term influences, especially concerning correlations of well-being with other variables where recall error weighs less severe (Powers et al. 1978).

For the discussion and interpretation of well-being developments, the reference system used for the corresponding assessments is important. When reports of current life satisfaction from different time points are used the reference systems can be different which leads to a mutatis mutandis-interpretation of the psychological concept of experienced utility. A stable reference system permits a ceteris paribus-interpretation that comes closer to the economic concept of decision utility (Easterlin 2001, 2002). The current, retrospective and prospective data thus can complement each other to address different questions. The significant association of past predictions and current recall of well-being, controlling for fixed personality traits and concurrent life satisfaction suggests that changes in well-being, in particular if expected or small, are not necessarily mirrored by changes in current reports but in ex-post updates of past assessments which offers an additional perspective on adaptation and response shift. In this case, current life satisfaction data can describe adequately the gross development of personal well-being (including changes in the personal, social or aspirational reference system) while intertemporal, especially retrospective, accounts can provide additional information about which life situation (e.g., now or a year ago) individuals net prefer. Not least, this distinction is relevant for policy questions about whether a particular development is rooted in externally objective living conditions or in the way the conditions are assessed internally. Overall, a case can be made for eliciting intertemporal and other, e.g., interpersonal comparisons, from survey participants directly (van Praag 2011). For policy analysis, a combination of different methodological approaches, including different well-being measures and based on decision as well as experienced utility would be preferable (Larsen and Fredrickson 1999; Loewenstein and Ubel 2008).

---

13 Kaiser and Vendrik (2020) stress that heterogeneous effects across life satisfaction scales and heterogeneity in the use of life satisfaction scales are likely and should be studied in addition to average effects.

14 Kahneman et al. (1997) provide a thorough discussion of the two concepts, including a distinction between current and retrospective assessments of experienced utility (see also Kahneman and Thaler 2006).
Funding Open Access funding enabled and organized by Projekt DEAL. The paper solely reflects the personal views of the author. The author has no financial or proprietary interests in any material discussed in this article. No funds, grants, or other support was received for conducting this study. The data is available to researchers from the German Socio-Economic Panel at the DIW Berlin; the code used in the empirical analysis is available from the author upon request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Baetschmann, G., Staub, K., & Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society Series A*, 178(3), 685–703. https://doi.org/10.1111/rssa.12090.

Baetschmann, G., Ballantyne, A., Staub, K., & Winkelmann, R. (2020). feologit: A new command for fitting fixed-effects ordered logit models. *Stata Journal*, 20(2), 253–275. https://doi.org/10.1177/1536867X20930984.

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91(2), 67–72. https://doi.org/10.1257/aer.91.2.67.

Bond, T., & Lang, K. (2019). The sad truth about happiness scales. *Journal of Political Economy*, 127(4), 1629–1640. https://doi.org/10.1086/701679.

Clark, A. (2018). Four decades of the economics of happiness: Where next? *Review of Income and Wealth*, 64(2), 245–269. https://doi.org/10.1111/roiw.12369.

Clark, A., & Georgellis, Y. (2013). Back to baseline in Britain: Adaptation in the British household panel survey. *Economica*, 80(319), 496–512. https://doi.org/10.1111/j.1468-0335.2011.00896.x.

Clark, A., Diener, E., Georgellis, Y., & Lucas, R. (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *Economic Journal*, 118(529), F222–F243. https://doi.org/10.1111/j.1468-0297.2008.02150.x.

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. https://doi.org/10.1257/jel.46.1.95.

Clark, A., D’Ambrosio, C., & Ghislandi, S. (2016). Adaptation to poverty in long-run panel data. *The Review of Economics and Statistics*, 98(3), 591–600. https://doi.org/10.1162/REST_a_00544.

D’Ambrosio, C., & Frick, J. (2012). Individual wellbeing in a dynamic perspective. *Economica*, 79(314), 284–302. https://doi.org/10.1111/j.1468-0335.2011.00896.x.

Deaton, A. (2012). The financial crisis and the well-being of Americans: 2011 OEP Hicks Lecture. *Oxford Economic Papers*, 64(1), 1–26. https://doi.org/10.1093/oep/gpr051.

Deaton, A., & Stone, A. (2013). Two happiness puzzles. *American Economic Review*, 103(3), 591–97. https://doi.org/10.1257/aer.103.3.591.

Deaton, A., & Stone, A. (2016). Understanding context effects for a measure of life evaluation: How responses matter. *Oxford Economic Papers*, 68(4), 861–870. https://doi.org/10.1093/oep/gpw022.

Di Tella, R., & MacCulloch, R. (2006). Some uses of happiness data in economics. *Journal of Economic Perspectives*, 20(1), 25–46. https://doi.org/10.1257/0895330067676526111.

Di Tella, R., Haisken-DeNew, J., & MacCulloch, R. (2010). Happiness adaptation to income and to status in an individual panel. *Journal of Economic Behavior & Organization*, 76(3), 834–852. https://doi.org/10.1016/j.jebo.2010.09.016.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, 29(1), 94–122. https://doi.org/10.1016/j.joep.2007.09.001.

Dubé, L., Jodoin, M., & Kairouz, S. (1998). On the cognitive basis of subjective well-being analysis: What do individuals have to say about it? *Canadian Journal of Behavioural Science*, 30(1), 1–13. https://doi.org/10.1037/h0087053.
Easterlin, R. (2001). Income and happiness: Towards a unified theory. *Economic Journal, 111*(473), 465–484. https://doi.org/10.1111/1468-0297.00646.

Easterlin, R. (2002). Is reported happiness five years ago comparable to present happiness? A cautionary note. *Journal of Happiness Studies, 3*(2), 193–198. https://doi.org/10.1023/A:1019608227505.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal, 114*(497), 641–659. https://doi.org/10.1111/j.1468-0297.2004.00235.x.

Frey, B., & Stutzer, A. (2002). What can economists learn from happiness research? *Journal of Economic Literature, 40*(2), 402–435. https://doi.org/10.1257/002205102320161320.

Frijters, P., Greenwell, H., Haissken-DeNew, J., & Shields, M. (2009). How well do individuals predict their future life satisfaction? Evidence from panel data following a nationwide exogenous shock. *Canadian Journal of Economics, 42*(4), 1326–1346. https://doi.org/10.1111/j.1540-5982.2009.01547.x.

Gandelman, N., & Hernandez-Murillo, R. (2009). The impact of inflation and unemployment on subjective personal and country evaluations. *Federal Reserve Bank of St Louis Review, 91*(3), 107–126. https://doi.org/10.20995/r.91.107-126.

Gibson, J., & Kim, B. (2010). Non-classical measurement error in long-term retrospective recall surveys. *Oxford Bulletin of Economics and Statistics, 72*(5), 687–695. https://doi.org/10.1111/j.1468-0084.2010.00599.x.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics, 19*(3), 403–420. https://doi.org/10.1016/s0167-6296(99)00037-5.

Groot, W., & Maassen van den Brink, H. (2000). Life-satisfaction and preference drift. *Social Indicators Research, 50*(2), 315–328. https://doi.org/10.1023/A:1007085500976.

Hagerty, M. (2003). Was life better in the “good old days”? Intertemporal judgments of life satisfaction. *Journal of Happiness Studies, 4*, 115–139. https://doi.org/10.1023/A:1024406800912.

Kahneman, D., & Thaler, R. (2006). Anomalies: Utility maximization and experienced utility. *Quarterly Journal of Economics, 112*(2), 375–405. https://doi.org/10.1162/00335537555235.

Kaiser, C. (2020). *How threatening are transformations of happiness scales to subjective wellbeing research?* ZBW - Leibniz Information Centre for Economics. https://hdl.handle.net/10419/226218.

Kaiser, C., & Vendrik, M. (2020). *How threatening are transformations of happiness scales to subjective wellbeing research?* IZA Discussion Papers 13905, Institute of Labor Economics (IZA). http://hdl.handle.net/10419/226218.

Kapteyn, A., van Praag, B., & van Herwaarden, F. (1978). Individual welfare functions and social reference spaces. *Economics Letters, 1*(2), 173–177. https://doi.org/10.1016/0165-1765(78)90057-5.

Krueger, A., & Schkade, D. (2008). The reliability of subjective well-being measures. *Journal of Public Economics, 92*(8–9), 1833–1845. https://doi.org/10.1016/j.jpubeco.2007.12.015.

Larsen, R., & Fredrickson, B. (1997). Back to Bentham? Explorations of experienced utility. *Behavior & Organization, 20*(1), 221–234. https://doi.org/10.1257/002205102320161320.

Larsen, R., & Fredrickson, B. (1999). Measurement issues in emotion research. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 40–60). New York: Russell Sage Foundation.

Loewenstein, G., & Ubel, P. (2008). Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics, 92*(8–9), 1795–1810. https://doi.org/10.1016/j.jpubeco.2007.12.011.

Manzonl, A., Luijkx, R., & Muffels, R. (2011). Explaining differences in labor market transitions between panel and life-course data in West Germany. *Quality & Quantity, 45*(2), 241–261.

McBride, M. (2010). Money, happiness, and aspirations: An experimental study. *Journal of Economic Behavior & Organization, 74*(3), 262–276. https://doi.org/10.1016/j.jebo.2010.03.002.

Odermatt, R., & Stutzer, A. (2019). (Mis-)predicted subjective well-being following life events. *Journal of Economic Literature, 37*(1), 425–483. https://doi.org/10.1093/jeea/jvy005.

Oishi, S., Schimmack, U., & Colcombe, S. (2003). The contextual and systematic nature of life satisfaction judgments. *Journal of Experimental Social Psychology, 39*, 232–247. https://doi.org/10.1016/S0022-1031(03)00016-7.

Pavot, W., & Diener, E. (1993). The affective and cognitive context of self-reported measures of subjective well-being. *Social Indicators Research, 28*, 1–20. https://doi.org/10.1007/BF01086714.

Powers, E., Goudy, W., & Keith, P. (1978). Congruence between panel and recall data in longitudinal research. *Public Opinion Quarterly, 42*(3), 380–389. https://doi.org/10.1086/268461.

Prati, A., & Senik, C. (2020). *Feeling good or feeling better?* IZA Discussion Papers 13166, Institute of Labor Economics (IZA). http://hdl.handle.net/10419/216478.
Schaffer, M. (2015). *GINIREG: Stata module for Gini regression*. https://EconPapers.repec.org/RePEc:boc:bocode:s457958.

Schröder, C., & Yitzhaki, S. (2017). Revisiting the evidence for cardinal treatment of ordinal variables. *European Economic Review*, 92(C), 337–358. https://doi.org/10.1016/j.euroecorev.2016.12.011.

Schwandt, H. (2016). Unmet aspirations as an explanation for the age u-shape in wellbeing. *Journal of Economic Behavior & Organization*, 122, 75–87. https://doi.org/10.1016/j.jebo.2015.11.011.

Schwartz, C. E., & Sprangers, M. A. (1999). Editorial. *Social Science & Medicine*, 48(11), 1505–1506. https://doi.org/10.1016/S0277-9536(99)00044-1.

Schwartz, C. E., & Sprangers, M. A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48(11), 1531–1548. https://doi.org/10.1016/S0277-9536(99)00047-7.

Schwarz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 61–84). New York: Russell Sage Foundation.

Senik, C. (2009). Direct evidence on income comparisons and their welfare effects. *Journal of Economic Behavior & Organization*, 72(1), 408–424. https://doi.org/10.1016/j.jebo.2009.04.019.

Smith, D., Loewenstein, G., Jepson, C., Jankovich, A., Feldman, H., & Ubel, P. (2008). Mispredicting and misremembering: Patients with renal failure overestimate improvements in quality of life after a kidney transplant. *Health Psychology*, 27(5), 653–658. https://doi.org/10.1037/a001264.7.

Schwartz, M. C. (2010). Do not throw out the baby with the bath water: Build on current approaches to realize conceptual clarity response to Ubel, Peeters and Smith. *Quality of Life Research*, 19(4), 477–479. https://doi.org/10.1007/s11136-010-9611-y.

Steffel, M., & Oppenheimer, D. (2009). Happy by what standard? The role of interpersonal and intrapersonal comparisons in ratings of happiness. *Social Indicators Research*, 92, 69–79. https://doi.org/10.1007/s11205-008-9289-5.

Stutzer, A., & Frey, B. S. (2010). Recent advances in the economics of individual subjective well-being. *Social Research: An International Quarterly*, 77(2), 679–714. https://www.muse.jhu.edu/article/528044.

Ubel, P., Peeters, Y., & Smith, D. (2010). Abandoning the language of response shift: A plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. *Quality of Life Research*, 19(4), 465–471. https://doi.org/10.1007/s11136-010-9592-x.

van Praag, B. (1971). The welfare function of income in belgium: An empirical investigation. *European Economic Review*, 2(3), 337–369. https://doi.org/10.1016/0014-2921(71)90045-6.

van Praag, B. (2011). Well-being inequality and reference groups: An agenda for new research. *Journal of Economic Inequality*, 9(1), 111–127. https://doi.org/10.1007/s10888-010-9127-2.

Wagner, G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel sStudy (SOEP) - scope, evolution and enhancements. *Schmollers Jahrbuch: Journal of Applied Social Science Studies Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 127(1), 139–169. http://hdl.handle.net/10419/150550.

Wunder, C. (2012). Does subjective well-being dynamically adjust to circumstances? *Economics Letters*, 117(3), 750–752. https://doi.org/10.1016/j.econlet.2012.08.024.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.