Feeling good is feeling better

This manuscript was submitted for consideration at Psychological Science

Can people remember their past happiness? We study four longitudinal surveys from the United States, France, United Kingdom and Germany, spanning from the 70s until today, where more than 60,000 people were asked questions about their current and past life satisfaction. We uncover systematic biases in recalled happiness: on average, people tend to overstate the improvement in their well-being over time and to understate their past happiness. But this aggregate figure hides a deep asymmetry: while happy people recall the evolution of their life to be better than it was, unhappy ones tend to exaggerate its worsening. It thus seems that feeling happy today implies feeling better than yesterday. This recall structure bears implications on motivated memory and learning, and can explain why happy people are more optimistic, perceive risks to be lower and are more open to new experiences.

life satisfaction | remembered happiness | memory bias | intra-personal comparisons

Statement of Relevance

We compare annual reports of life satisfaction and their retrospective reconstruction using over 250,000 interviews from different countries and epochs. On average, people declare to be happier than in the past, although this is not consistent with their annual reports. While happy people overstate the positive evolution of their life, unhappy ones tend to exaggerate its worsening. These findings shed light on the interplay between happiness levels and variations, and help understanding differences in behavioural traits, such as optimism, attitudes to risk and to new experience.

When maximizing well-being (or happiness, or utility), people have to choose between several alternatives, the value of which they often evaluate, when possible, based on their memory of the happiness they derived in past similar experience. Because expectations are partly based on memory, remembered happiness is of importance for decision-making and memory biases create misleading pieces of information. So, how reliable is the memory of past happiness?

Previous research has focused on people's capacity to accurately recall the impact of a specific event on their happiness, like electoral results (Wilson et al., 2003; Kaplan et al., 2016) or medical treatments (Smith et al., 2008). This literature has documented structured biases in recalled happiness, but to date little is known about how people recall their previous overall happiness. However, it is often impossible to think about an event in a vacuum, and people's memories are typically governed by the general context of a period of their past life. The question of how people recall the hedonic quality of previous time episodes is not entirely new, and some studies have explored the structure of recalled happiness over a horizon measured in minutes or days (Kahneman et al., 1993; Fredrickson and Kahneman, 1993; Kemp et al., 2008; Ganzach and Yaor, 2019; Strijbosch et al., 2019). Yet, socially-relevant decisions are often based on the happiness attached to longer time spans, in the order of months or even years.

Measuring memory accuracy over long time spans is also a methodological question. From its inception (Bartlett, 1932) - and until today (Chew et al., 2019; Saucet and Villeval, 2019; Zimmermann, 2020) - the vast majority of empirical evidence on long-term memory has been based on experiments where subjects are asked to recall some information they previously learned or reported within the experiment. In spite of their many merits, experimental methods suffer from some important drawbacks: they take place over a relatively short amount of time, thus not allowing to assess long-term patterns, and often involve relatively small samples, thus hindering the identification of within-population heterogeneity. To make progress, this paper uses data from three publicly available national surveys and one commercial survey. It is composed of four studies, which take the opportunity of about 260,000 self-assessed life satisfaction judgements from more than 60,000 individuals along several years. By exploiting the longitudinal dimension of the surveys, we can compare people's actual life satisfaction with their retrospective evaluation. The studies complement each other in offering a rich picture on the structure of recalled happiness over years, as they ask the question about past happiness in different ways and over different time horizons.*

What is meant by "happiness" matters, too. Psychologists have long known that current happiness plays a role in misremembering (Blaney, 1986; Fiedler and Hütter, 2013) and that a memory-experience gap in happiness exists (Wilson et al., 2003; Smith et al., 2008). However, the main focus has been on affective states (emotional arousal and mood), rather than on cognitive evaluations (satisfaction). Neglecting this difference leaves an incomplete picture of the relationship between autobiographical memory and happiness, and may lead to inaccurate predictions, as different attributes of happiness can relate to different recall dynamics. For instance, memory biases follow opposite patterns whether the recollected feelings are either the general mood or some episode-related emotions (Kaplan et al., 2016), and specific emotions of the same valence can follow

*The term "memory" is very rich and complex. In the present study, we will use it uniquely to refer to self-referential long-term explicit memory. We refer to "biases" without any negative connotation, and refer to them as cases where people's current knowledge and beliefs distort their memory of the past (Schacter et al., 2003). The use of the term is compatible with the classic notion of bias in behavioral economics, viz. systematic errors of judgment which occur under specific conditions (Tversky and Kahneman, 1974).

Table 1. Summary statistics

	mean	sd	min	max
Study 1 ($N = 121,616$)				
life satisfaction	7.15	1.71	0	10
more satisfied than 10 years ago	0.52	0.27	0	1
about the same as 10 years ago	0.32	0.26	0	1
less satisfied than 10 years ago	0.16	0.06	0	1
Study 2 ($N = 111,894$)				
life satisfaction	5.19	1.25	1	7
more satisfied than a year ago	0.25	0.44	0	1
about the same as a year ago	0.58	0.49	0	1
less satisfied than a year ago	0.16	0.36	0	1
Study 3 ($N = 18,589$)				
life satisfaction	6.48	1.73	0	10
more satisfied than a year ago	0.30	0.45	0	1
about the same as a year ago	0.45	0.43	0	1
less satisfied than a year ago	0.25	0.50	0	1
Study 4 ($N \approx 10,000$)				
life satisfaction (1971-76)	6.50	2.03	0	10
life satisfaction (2001-06)	6.77	1.79	0	10
life satisfaction 5 y. ago (1971-76)	5.60	2.60	0	10
life satisfaction 5 y. ago (2001-06)	6.20	2.21	0	10

Note: Study 1 is based on German Socio-Economic Panel (SOEP), waves 23 to 33, 121,616 individual-year observations, from 11,056 respondents. Row 2: Life satisfaction: "How satisfied are you with your life, all things considered? Scale: 0-10". Rows 3-5: "Which of the nine pictures best represents how satisfied you have been with your personal living situation over the last 10 years, so from around 2006 to today?", question asked in wave 33 (2016). Possible answers: see fig. 1. Study 2 is based the British Household Panel Survey (BHPS), waves 6-10 and 12-18 (1997-2001 and 2002-2009), 111,894 individual-year observations from 20,269 respondents. Row 7: "How dissatisfied or satisfied are you with your life overall?" Scale: 1 "not satisfied at all" to 7 "completely satisfied". Rows 8-10: "Would you say that you are more satisfied with life, less satisfied or feel about the same as you did a year ago?". Possible answers: "more satisfied", "less satisfied", "about the same", "don't know". Study 3 is based on the French Consumer Confidence Survey (CAMME), 11 quarterly waves, from June 2016 to December 2018; 18,589 respondents. Row 12: "Overall, how satisfied are you with your current living situation on a scale from 0 to 10?". Rows 13-15: "When you think about last year, where would you place yourself on a scale from 0 to 10?". Study 4 is based on Gallup Poll Social Series, waves 1971, 1972, 1974, 1976 and waves 2001, 2002, 2003, 2004, 2005, 2006. Rows 17-18: "Please think about a picture of a ladder. Suppose that the top of the ladder represents the best possible life for you, and the bottom represents the worst possible life for you. If the top step is "10" and the bottom step is "0", on which step of the ladder do you feel you personally stand at the present time?". Rows 19-20 "On which step would you say you stood five years ago?". Column 1: mean value, column 2: standard deviation. The standard error of the survey-weight estimated mean is $\leq$ 0.01 for all means except row 1 ($se = 0.06$).

different recall patterns (Levine et al., 2021). We focus on measures of life satisfaction, a construct which captures people's general cognitive and hedonic evaluation of their life, as opposed to momentary evaluations (Kahneman and Krueger, 2006).

Despite life satisfaction data have been gaining considerable interest among social scientists (Clark, 2018), recalled life satisfaction has received little attention so far. Some exceptions are the classic contributions by Cantril (1965) and Easterlin (2001), and the small literature dedicated to hedonic adaptation (Clark, 1999; Lyubomirsky, 2011), but none of them analyzes the discrepancy between reported and recalled well-being. A couple of recent papers (Köke and Perino, 2017; Kaiser, 2020) noticed some inconsistencies between actual and recalled well-being in one of the surveys analyzed herein, but they did not attempt to explain the recall structure. A few methodological papers have used recalled subjective well-being to assess the test-retest reliability of life satisfaction scores (Atkinson, 1982; Michalos and Kahlke, 2010). They find that recalled evolution is imperfect but statistically consistent with the observed variations. Closer to our paper, a small series of studies has uncovered a general tendency of people to report an upward trajectory in life satisfaction ratings, which often implies over-estimating the improvement in their actual life satisfaction as compared to its past level (Busseri et al., 2009; Busseri and Samani, 2019). Other research added the observation that this discrepancy between actual and recalled evolution of life satisfaction is especially prevalent among the younger (Lachman et al., 2008; Busseri and Samani, 2019). Overall, they propose that this bias is consistent with a self-motivation device, whereby people are trying both to improve their own dynamic self-image and to create optimistic beliefs, an interpretation that is close to the theory of motivated beliefs (Bénabou and Tirole, 2002). Here, we enquire about the structure of these biases, and uncover an asymmetry that is not explained by socio-demographic variables.

We present four studies, that complement each other in the theoretical development and empirical assessment of a model of recalled happiness, which generates testable predictions. In its general form, we can describe the model as follows: people who are relatively satisfied with their life at time t tend to recall its long-term evolution more as a stable progression than as a fluctuation (study 1), to recall its variation to be better than it was (study 2) and its past level to be worse than it was (studies 3 and 4). Conversely, people who are relatively unsatisfied with their life tend to recall a more negative past evolution and exaggerate their past happiness. Our recall model contributes to the literature that investigates how retrospective reports of feelings are shifted in the direction of the current set of beliefs (Ross, 1989; Robinson and Clore, 2002; Levine et al., 2001; Kaplan et al., 2016).

Results

Most people feel that they are happier than they used to be in the past. In Studies 1, 2, 3 and 4 we analyze data from four nationally representative surveys which include questions on recalled happiness. Table 1 presents descriptive statistics for all studies. Average life satisfaction (LS) is relatively high in each sample. Most people (52%) report to be happier than 10 years ago and people tend to rate their current life satisfaction to be higher than five years ago. Many people report to be as satisfied as in the previous year, but the reported changes are asymmetrical: the share of respondents reporting a positive evolution is higher than the proportion reporting a negative evolution. This positive picture has been already illustrated before us (Cantril, 1965; Lachman et al., 2008; Busseri et al., 2009), in spite of the accumulated empirical evidence that life satisfaction does not steadily increase over the life cycle.[†]

[†] The average observed pattern is U-shaped, with a ditch around 50 (Blanchflower and Oswald, 2008).

Study 1. In study 1, we use the German Socio-Economic Panel (SOEP), one of the longest and most comprehensive panel surveys in the world. A representative sample of German residents is interviewed every year and in each wave of the survey respondents are asked to report on a scale from 0 to 10 "How satisfied are you with your life, all things considered? [0-10]". In 2016 (wave 33), they were asked the following question: "Which of the nine pictures best represents how satisfied you have been with your personal living situation over the last 10 years, so from around 2006 to today?". By comparing their answers to the chronicle of their yearly reports of satisfaction, we are able to check whether their subjective reconstructions are consistent.

People can, to some extent, recall their past happiness. Figure 1 illustrates conditional average observed patterns and compares them to the schematic representation that was chosen by respondents. It turns out that reported and observed patterns are quite close. When considering the difficulty of the cognitive task at stake - namely, recalling the evolution of a subjective state over a period as long as ten years - the consistency observed in Figure 1 is not trivial. It solidly corroborates the idea that people can actually recall - to some extent - their past happiness. It also offers supporting evidence for the stability of the satisfaction scale, a pivotal assumption needed for any longitudinal study on subjective well-being. Yet, some caution is needed in the interpretation of these results. Fig.1 neither proves that people use *exactly* the same scale over time (this is unlikely to be the case), nor that people recall their past happiness *perfectly* (as we will show in what follows). Nevertheless, it corroborates the hypothesis that recalled happiness is a meaningful construct, evaluated on a scale that is comparable over time.

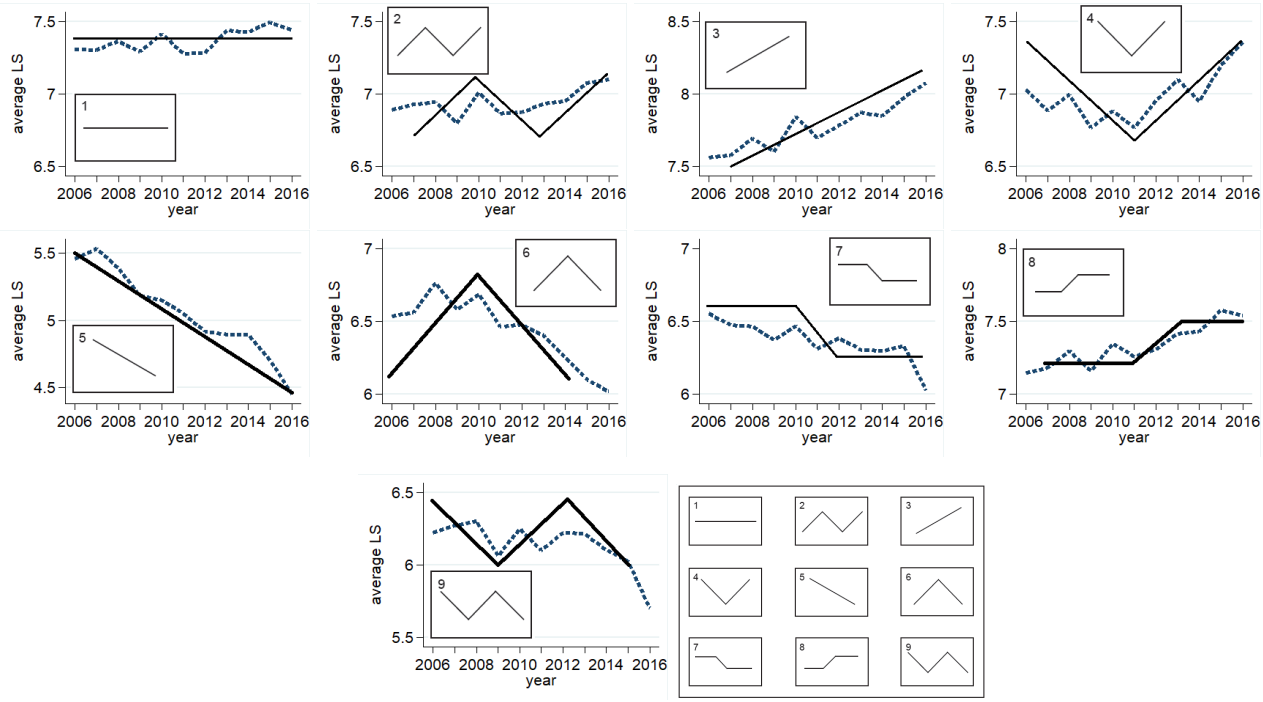


Fig. 1. Observed evolution of life satisfaction, conditional on the recalled pattern.

Note: Source: SOEP, waves 23-33 (2006-2016), 121, 616 individual-year points. Pictures 1-9: Observed long-run evolution of life satisfaction as reported yearly by each individual over 10 years ("How satisfied are you with your life, all things considered?") and declared long-run evolution ("Which of the nine pictures best represents how satisfied you have been with your personal living situation over the last 10 years, so from around 2006 to today?"). Picture 10: Show card from SOEP.

Figure 1 also offers some insights on the relationship between average current satisfaction and its recalled evolution. Consider pictures 2 (up-down-up), 3 (up-up-up) and 8 (flat-up-flat). More than half of the respondents describe their past by mean of one of these pictures (see SI Appendix, Table S1 for response frequencies). The actual longitudinal evolution is very similar, so that the average individual who chose picture 8, could have chosen equally picture 2 or picture 3 to describe their past. However, an important difference appears: people who recall different patterns are on different satisfaction levels. Those who are relatively happier ($LS = 8.1$) describe their past as a constant progression (picture 3), people in the middle ($LS = 7.5$) describe it as a weakly monotonic improvement (picture 8) and people who are relatively less happy ($LS = 7.1$) reconstruct a non-monotonic improvement (picture 2). This difference suggests some influence flowing from the current level of satisfaction to the recollection of past satisfaction. Does feeling good implies feeling better?

To examine the pattern in misrecalling, we restrict the sample to people who report similar levels of life satisfaction ten years apart, in 2006 and 2016 (N=5,993).[‡] A simple analysis shows that people who report positive patterns tend to be

[‡]To preserve a sufficiently large sample, we allow for a 1-point-discrepancy, so that $LS_{it} \approx LS_{it-10}$ if $LS_{it} = LS_{it-10} \pm 1$. If we adopt a strict definition and consider only cases when

significantly more satisfied than people who report negative patterns. On average, among people who display similar levels of happiness in 2006 and 2016, those who recall a steady positive increase (picture 3) report a satisfaction level higher than 8 at both points in time; those who recall a non-linear increasing pattern (pictures 2 and 8) report $LS \approx 7.5$; those who recall a non-linear decreasing pattern (pictures 7 and 9) report $LS \approx 6.5$; finally, people who recall a steady negative evolution (picture 5) are much less satisfied than the rest of the population in 2006 and 2016, $LS \approx 5$ (see SI Appendix Figure S1 for a visual illustration). In sum, among people who report inaccurate patterns, those who are relatively happier tend to recall a steadier positive evolution in their lives.

Study 2. While data from the German national survey analyzed in Study 1 provide confidence in people's capacity to reconstruct patterns of their past happiness, they also point toward some relationship between levels and recalled variations, which cannot be tested without more granular data. To provide a more accurate picture, we use publicly available data from the British Household Panel Survey (BHPS), a panel survey of a representative sample of British residents. From 1997 to 2009 (except in 2002) respondents were asked about their general life satisfaction, both as level ("How dissatisfied or satisfied are you with your life overall?" [1-7]]) and as compared to the previous year ("Would you say that you are more satisfied with life, less satisfied or feel about the same as you did a year ago?" ["more satisfied", "less satisfied", "about the same", "don't know"]. By comparing individual answers over time, we are able to infer discrepancies between the observed (computed by the researcher, based on annual LS) and the reported (declared by the respondent, based on their memories) change in life satisfaction. We study 111,894 individual-year observations from 20,269 respondents.

We start our analysis by cross-tabulating reported and observed changes in life satisfaction. We can distinguish 9 cases, displayed in Table 2. The table confirms that people are, to some extent, able to recall the evolution of their well-being. If reported changes were uncorrelated to observed changes, the nine cells of the table would display similar percentages. Instead, when we observe an increase in the life satisfaction score (26.3% of the total), people report a positive evolution three times more often than a negative one; whereas when we observe a decrease in the life satisfaction score (27.7% of the total), people report a negative evolution twice as often as a positive one. When we observe no change (46.0% of the total), more than half of the sample reports to be as satisfied as in the previous year. Nevertheless, discrepancies between observed and reported changes are far from rare. In what follows, we analyze these discrepancies.

Table 2. Cases of reporting behaviors. $N = 111,894$

		Observed			
		−	=	+	
Reported	−	8.3%	5.4%	2.8%	16.5%
	=	15.0%	28.5%	14.7%	58.2%
	+	4.4%	12.1%	8.8%	25.3%
		27.7%	46.0%	26.3%	100%

Note: Source: BHPS, waves 6-18 (1997-2009), 111,894 individual-year points. Reported change: "Would you say that you are more satisfied with life, less satisfied or feel about the same as you did a year ago?" ("more satisfied", "less satisfied", "about the same", "don't know"). Observed change: difference between current life satisfaction (LS) and LS reported by the same individual last year. 8.3% answers report a negative change in life satisfaction which is consistent with what we observe. Overall, 27.7% of observed changes are negative and 16.5% of reported changes are negative. See the Methods Section.

Happy people overstate the improvement in their life satisfaction; conversely, unhappy people understate it. For the sake of guidance over our analysis, we classify reporting behaviors into four broad types. The same individual can display different behaviors along the panel, but he/she displays only one reporting behavior per year.

- *correct-reporting*: Report an evolution of life satisfaction that is consistent with what we observe.
- *over-reporting*: Report a positive change in life satisfaction that we do not observe.
- *insensitive-reporting*: Report no change in life satisfaction but we observe one.
- *under-reporting*: Report a negative change in life satisfaction that we do not observe.

Overall, correct-reporters represent a little less than half of the sample, and along the panel over 90% of the respondents incorrectly report at least once. To explore the relationship between reporting behaviors and the current level of satisfaction, we start with a simple analysis of the first cross-section of the BHPS (wave 7) that includes the question "Would you say that you are more satisfied with life, less satisfied or feel about the same as you did a year ago?". Two opposite patterns appear (see fig.2): over-reporters score significantly higher in life satisfaction than the rest of the population ($t(7928) = 9.583$, $p < 0.001$); conversely, under-reporters' life satisfaction is significantly lower than for the rest of the population ($t(7928) = -16.70$, $p < 0.001$). There is no significant difference between correct-reporters and insensitive-reporters ($t(5864) = 0.20$, $p = 0.838$). The uncovered patterns do not seem to be driven by other observable traits, like income or age. To test this, we carry on a cross-sectional regression by standard OLS (see SI Appendix Table S2) and compute the vector of residuals, which is orthogonal to the explanatory variables of the model, by construction. We can interpret this vector as the individual-specific residual life satisfaction, conditional on the observables. Figure 2 illustrates the average residual satisfaction by group of reporting behavior.

$LS_{it} = LS_{it-10}$ (2,676 observations), results are qualitatively similar but confidence intervals are sensibly larger.

The uncovered patterns are the same as before: mean residual life satisfaction is higher for over-reporters ($t(7928) = 9.542$, $p < 0.001$) and lower for under-reporters ($t(7928) = -15.07$, $p < 0.001$), while it is statistically equal to zero for the rest of the sample ($t(5783) = 0.87$, $p = 0.382$).

It is worth noting two points. Firstly, the asymmetry in Figure 2 is not mechanical. Quite the opposite, the dynamic movement along the life satisfaction scale would predict the reverse: people who report a relatively high satisfaction in wave t are also the ones for whom we are more likely to observe a positive evolution. And if a positive evolution is observed, over-reporting cannot happen since people reporting a positive change would fall into the category of correct-reporters. Thus, *a priori*, one would expect over-reporters to be relatively less happy, and under-reporters to be relatively happier. Secondly, this asymmetry does not follow from individual variations in material conditions (see also SI Appendix Figure S2), and it is not driven by a particular age group (see also SI Appendix Figure S4).

We now move on to the dynamic panel dimension. When we sketch the frequency of under- and over-reporting conditional on life satisfaction, two clear patterns appear (Figure 3a). On the one hand, the share of people declaring a *positive* change that we do not observe is significantly higher among satisfied respondents than among unsatisfied ones. On the other hand, the share of people declaring a negative change that we do not observe is significantly lower among satisfied respondents than among unsatisfied ones. The share of people reporting no change or a change which is consistent with what we observe in the panel is constant across levels of life satisfaction, except for an increase on the upper boundary. The figure also illustrates the the quantitative importance of under and over-reporting: overall, they concern about one fourth of the population.

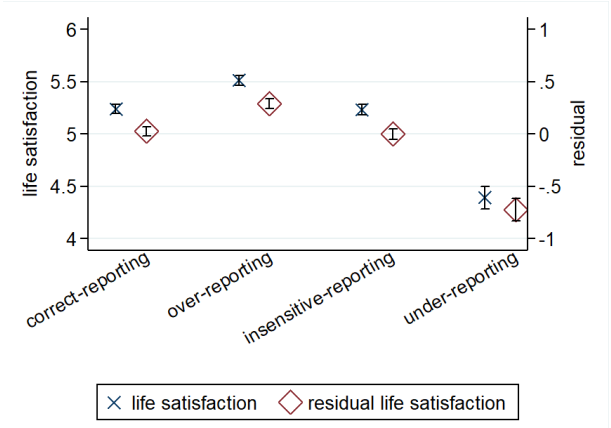
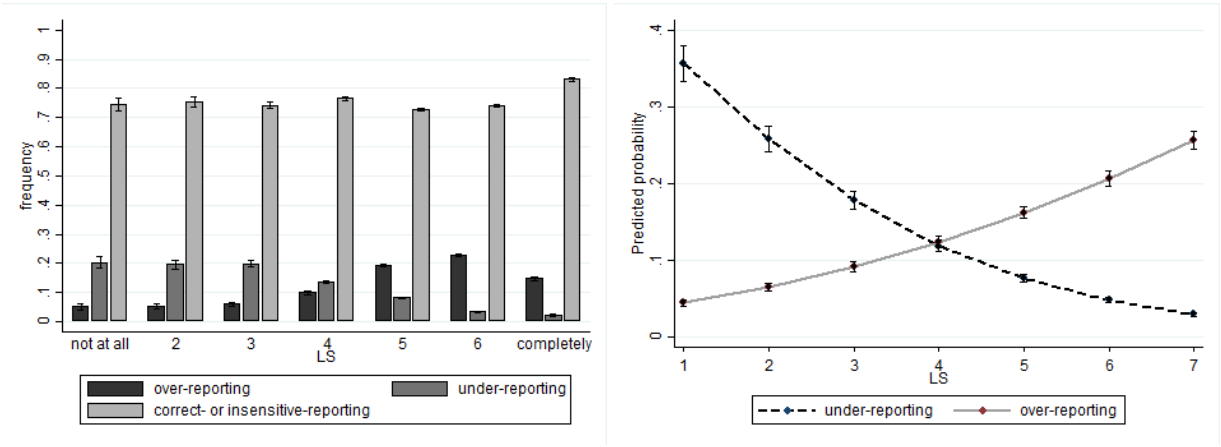


Fig. 2. Conditional and unconditional life satisfaction, by reporting behavior, wave 7

Note: Source: BHPS, waves 6 and 7 (1997-1998). Left-hand axis: mean LS. Correct-reporters are people who report a change in subjective well-being that is consistent with what is observed in waves 6 and 7. Over-reporters (resp. under-reporters) are people who report a positive (resp. negative) change in subjective well-being that is not consistent with LS in waves 6 and 7. Correct-reporters declare a level of life satisfaction of 5.2. The level of LS of people who over-report and under-report is, respectively 5.5 and 4.4. Right-hand axis: mean LS conditional on sex, age, age squared, log(household income), log(individual income), education level, number of children, marital status, job status, ethnicity and demographic change dummies. Everything else equal, people who over-report the change in their LS declare a level of LS 0.3 scale-points higher than correct reporters and 1 scale-point higher than under-reporters. Vertical bars are 95% confidence intervals.

To provide an estimate of how likely it is that happy and unhappy people recall their past differently, we use regression analysis (see SI Appendix "Supplementary Methods" for the conceptual framework and underlying hypotheses). We estimate the conditional probability to observe a given reporting behavior by Multinomial Logit, where we set correct-reporting as the base category. Results substantially confirm the previous analysis: the higher the life satisfaction level today, the higher the probability to over-report and the lower the probability to under-report. For instance, the average individual has respectively 8% chance to over-report and 17% chance to under-report if she declares $LS = 3$ at the moment of the interview. If instead her life satisfaction is high (LS=6), she has 18% chance to over-report and 5% chance to under-report. Regardless of her satisfaction level, her chance to correct-report (48% if $LS = 6$ and 44% if LS =3) or to insensitive-report (30% if $LS = 6$ and 31% if LS =3) are very similar. The estimated probabilities to over- and under-report are graphically displayed in Figure 3b. Regression estimates and estimated probabilities are listed in SI Appendix Tables S4 and S5. Results are qualitatively the same when we restrict our focus to the exemplary case of people who declare the same satisfaction for two consecutive years, but recall a change (see SI Appendix Table S4 and Figure S3).

Could this phenomenon simply reflects some measurement errors due to the life satisfaction scale? Scale boundaries cannot explain the overall patterns, which are quite stable along the life satisfaction scale, with no spike at the boundaries (in Figure 3a we observe a small fall, if any, in the share of over-reporting at the top of the scale). Moreover, regression results hold when we use binary indicators for the different levels of satisfaction, to single out effects at the upper- and lower-bound (SI Appendix Table S3). What about small stochastic changes that are not captured by the scale? Indeed, people may report a change in life satisfaction because latent satisfaction - which we can consider as a continuous variable - varies, but this variation is not captured by the satisfaction scale - which is discrete. This measurement error could explain over- and under-reporting if small changes occur within the range of the same life satisfaction scale point: people would rightfully report a (small) change, but we would wrongly infer misreporting. However, if we model the marginal change as a random draw from a symmetrical



(a) Observed shares of over- and under-reporting

(b) Predicted probabilities of over- and under-reporting

Fig. 3. Mis-reporting conditional on current life satisfaction

Note: Source: BHPS waves 6-18 (1997-2009). Fig. (a): distribution of reporting behavior depending of the level of LS. Among people who report a level of LS=6, 4% under-report the evolution of their LS, 23% over-report it and the remaining 86% are correct- or insensitive-reporters. Fig. (b): predicted probability of over and under-reporting based on a first stage regression of the probability to report a given outcome {more satisfied, about the same, less satisfied} over current life satisfaction and other covariates such as sex, age, income, education level, number of children, marital status, job status and ethnicity and demographic change. For a person whose level of satisfaction is 6, the probability to under-report is 5%, versus 20% to over-report. To avoid information burden, predictions for correct- and insensitive-reporting are not showed in the graphic, but only in SI Appendix (Table S5).

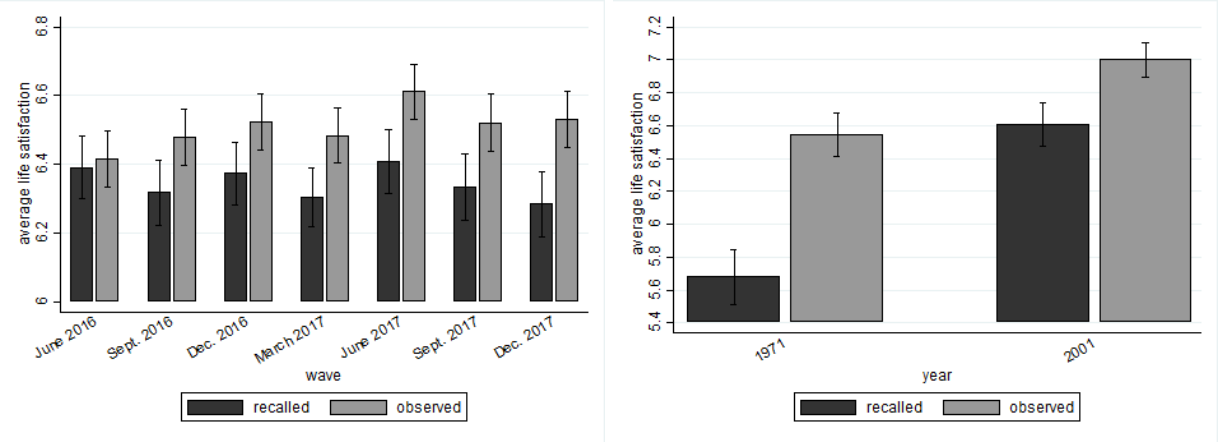
distribution centered on the latent value, we should be equally likely to observe under- and over-reporting. This is not the case since, on average, people tend to overstate the improvement in their life satisfaction. We examine this tendency in Study 3.

Study 3. Study 2 reveals that happy people tend to overstate the time improvement in their life satisfaction, whereas unhappy people tend to overstate the decline in their life satisfaction. Because the average level of life satisfaction in the population is fairly high (7 on a 0-10 scale, or 5 on a 1-7 scale), the overestimation bias tends to dominate the underestimation bias. Indeed, a quick glance at Table 2 tells us that cases of over-reporting are much more common than cases of under-reporting (16.5% vs 8.2%). But does people's tendency to feel better imply a tendency to downplay their past? If a person who reports the evolution of her life satisfaction from a previous period until now is comparing her current satisfaction - which is perfect information - with her past satisfaction - which is mediated by memory - then overstating the evolution of well-being implies understating the past level of well-being. To test this prediction, we would need to elicit the recalled *level* (instead of the variation) of well-being. Study 3 makes use of a survey that contains exactly this question. Moreover, as people are interviewed only once in that survey, we can be sure that they are trying to remember their past happiness level, rather than the numerical response they gave in the previous interview.

People tend to understate their past happiness and to overstate the improvement in their happiness over time. In Study 3, we use quarterly data drawn from the Consumer Confidence Survey (CAMME), run by the French national statistical office. On top of a question about current life satisfaction ("Overall, how satisfied are you with your current living situation?" [0-10]), the survey also includes a question about the level of satisfaction one year ago ("When you think about last year, where would you place yourself on a scale from 0 to 10?"). Therefore, this survey allows to observe for the same individual the current level of satisfaction and the remembered level of satisfaction one year ago. People are not interviewed twice, so that the actual past satisfaction of a given person (as experienced last year) is not available. However, insofar as each sample is representative of the same population, we can compare aggregate data on observed and recalled life satisfaction at the national level. Loosely speaking, we can consider France as an individual and compare its recalled and observed levels of well-being at a twelve-month gap. Graph 4a contrasts recalled and observed happiness over seven quarters.

Consistently with our predictions, at each quarter, the average reported well-being at a point of time is higher than the average well-being recalled for the same period. Differences are statistically significant at 10% or 5% for all periods except the first one. On average, people tend to downplay their past happiness.

Study 4. Across studies 1-3 we provided evidence of a specific structure of recalled happiness based on European data from the last three decades. Yet, cultural and historical factors can influence retrospective preferences so that it is unclear to what extent our findings are generalizable. Study 4 explores this avenue by using aggregate data collected by Gallup in the early 70s and in the early 2000s. In 1971 and 1976, as well as thirty years later, in 2001 and 2006, the Gallup team interviewed a representative sample of the American population and asked them the following questions "Please think about a picture of a ladder. Suppose that the top of the ladder represents the best possible life for you, and the bottom represents the worst possible life for you. If the top step is "10" and the bottom step is "0", on which step of the ladder do you feel you personally



(a) France (2016-17)

(b) US (1971, 2001)

Fig. 4. Average observed and recalled life satisfaction

Note: Fig. (a): source: CAMME waves 1-11 (June 2016 - December 2018); 18,589 respondents. Observed LS: "Overall, how satisfied are you with your current living situation on a scale from 0 to 10?". Recalled LS: "When you think about last year, where would you place yourself on a scale from 0 to 10?". The gray bars refer to the average current LS reported in period t. The black bars refer to the average recalled LS reported one year later, in $t+1$. In France, the average life satisfaction in June 2017 is 6.6, while the average recalled life satisfaction for the same period is 6.4. Vertical bars are 95% confidence intervals. Fig. (b): source: Gallup Poll Social Series, 4 waves, 1971, 1976, 2001 and 2006. About 4,000 respondents. Observed LS: "Please think about a picture of a ladder. Suppose that the top of the ladder represents the best possible life for you, and the bottom represents the worst possible life for you. If the top step is "10" and the bottom step is "0", on which step of the ladder do you feel you personally stand at the present time?". Recalled LS: "On which step would you say you stood five years ago?". In the United States, the average life satisfaction in January 2001 is 7.0, while the average recalled life satisfaction for the same period is 6.6. Vertical bars are 95% confidence intervals.

stand at the present time?" and "On which step would you say you stood five years ago?". We adopt a similar analytical method to that of Study 3 and compare current and recalled life satisfaction for the same year. Results, displayed in figure 4b, essentially replicate the ones from France: on either side of the Atlantic Ocean, 50 years ago as well as 20 years ago, people tend to recall their past happiness to be lower than it actually was ($t(969)$=-8.94, $p < 0.001$; $t(989)$=-5.01, $p < 0.001$).

Discussion

People's remembered well-being seems to be influenced by their current level of life satisfaction. While happy people tend to overstate the positive evolution of their life satisfaction, unhappy ones tend to exaggerate its worsening. The asymmetric biases that we uncovered cannot be explained by the limits of the measurement tool that we use, i.e. happiness scales. Rather, they seem to derive from purely behavioral factors.

Our findings clearly point to the confusion between levels and variations of happiness. Existing behavioral theories of subjective well-being have abundantly showed that differences and gaps often matter more for individual happiness than absolute levels (of income for instance). Here, we document another type of confusion between levels and variations. We uncover the influence that flows from a person's level of happiness to their remembered past happiness, and the dynamic evolution thereof.

These findings contribute to the existing literature both on autobiographical memory and on subjective well-being in different ways. The (partial) correspondence between recalled happiness trajectories and observed happiness trajectories provides an unprecedented test that rules out the threat of full adaptation of the life satisfactions scale.[§] The documented tendency to overstate the improvements in happiness supports the hypothesis that people prefer improving sequences not only in prospective judgments (Frank and Hutchens, 1993; Caplin and Leahy, 2001; Loewenstein and Sicherman, 1991; Loewenstein, 2006; Senik, 2008), but also in retrospective judgements. It also complements recent experimental evidence on motivated memory for self-relevant outcomes (Saucet and Villeval, 2019; Zimmermann, 2020). Thanks to the study of a long time span, it shows that most people tend to recall an improving happiness profile but thus tend to downplay - rather than outplay - their past well-being. These results invite reconsideration of the hypothesis that recalling positive self-relevant outcomes is ego-enhancing: comparisons with an inferior past may be beneficial (Tversky and Griffin, 1991; Wilson and Ross, 2000, 2001). Finally, the observed asymmetry in recall patterns between happy and unhappy suggests the existence of some heterogeneity in coping mechanisms and *ex-post* rationalization, that might remain veiled in small samples. This misperception structure can have important behavioral implications. In particular, it could create a self-reinforcing divergence between happy and unhappy people. Indeed, if happy people tend to think that they were less happy in the past, they will tend to be less conservative and more open to innovation. By contrast, if unhappy people believe that they used to be happier in the past, they could tend to be backward looking, hence be more attached to the status quo ante. This could constitute a new and additional explanation of the reason why happy people are more optimistic (Foster et al., 2012), perceive risks to be lower (Johnson and Tversky, 1983)

[§] If the scale completely shifted over time, the sole way to observe figure 1 would be that recall errors and memory shifts compensate each other. This hypothesis is not impossible, but at least very unlikely.

and are more open to new experiences (Furnham and Petrides, 2003), while unhappy people are more pessimistic, reluctant to change and perceive higher risk.

This paper is limited by the availability of data on recalled happiness and some important issues should be addressed in future research. Firstly, despite evidence that the recalled hedonic quality of an experience underlies decision making (Kahneman et al., 1993), this paper does not directly study the latter. Its natural extension is to look at how the way people recall their life narratives influences their behavior. This avenue for research belongs to a growing body of evidence that documents the predictive power of reported happiness for subsequent behavior, for example when it comes to productivity (De Neve and Oswald, 2012; Oswald et al., 2015; Bellet et al., 2019), and for subsequent health outcomes, for example morbidity, mortality and healthcare utilization (Goel et al., 2018; Rosella et al., 2019).

Secondly, our studies are based on the standard assumption that scales are invariant. This seems to be a reasonable approximation *on average* (see fig. 1), but alternative avenues are possible. Theoretically, if happy and unhappy people use their scale differently (as recently conjectured by Kaiser (2020)), and if the population *actually* gets happier over time, this rescaling process could explain the patterns we observe and help solve the long-standing Easterlin puzzle. As we show in this paper, data on recalled happiness can help jointly testing some hypotheses on the use of the satisfaction scale and on the recall process. We believe that the use of transition ratings in national surveys (e.g. "Overall, how dissatisfied or satisfied are you as compared to last year? [-5, +5]"), which are common in clinical studies but virtually absent in subjective well-being research, could greatly illuminate this path.

Lastly, although our results are compatible with the functional role of different misrecalling patterns, we cannot directly assess functional aspects, since this would introduce a circular element (people feel better because they feel good because they feel better and so on). Nevertheless, as a first exploratory step in this direction, we analysed twelve mental health outcomes which are measured in the British panel survey (GHQ12). It turns out that the individual propensity to revise one's own past in a positive or negative way correlates with all dimensions of mental health, even when controlling for a given life satisfaction level (SI, Tables S6 and S7). The higher the propensity to over-report, the better the mental health outcomes; the higher the propensity to under-report, the lower the mental health outcomes. Individual propensities to over- and under-report also predict some aspects of *future* mental health, like general happiness and sense of playing a useful role (SI, Tables S8). In future research, it will be important to understand the role of recall patterns in shaping emotions, attitudes and beliefs.

Materials and Methods

Details of the measured variables and mathematical modeling can be found in SI Appendix, section "Supplementary Methods".

Study 1. *Sample and procedures.* Since 1984, the German Socio-Economic Panel (SOEP) has been interviewing face-to-face a representative sample of the German population on a variety of topics, including subjective well-being. The sample is a stratified clustered design with 125 Primary Sampling Units. It has included the states of the former German Democratic Republic since 1990. SOEP is funded by the German Federal Government and different agencies. Data are available free of charge through the SOEP Research Data Center, upon agreement with the German Institute for Economic Research (DIW Berlin). Goebel et al. (2019) offers an up-to-date comprehensive overview of the dataset. We exclude the attrition sample and construct a balanced panel of life satisfaction data from waves 23 to 33, which offers 121,616 individual-year observations, from 11,056 respondents (53% women, average age 52).

Study 2. *Sample and procedures.* The British Household Panel Survey (BHPS) is a panel survey of a representative sample of British residents which was run between 1991 and 2009. Interviews took place face-to-face, annually and lasted approximately 30-40 minutes. The sample was a stratified clustered design with 250 Primary Sampling Units. The survey was representative of Wales and Scotland since 1999, and included Northern Ireland since 2001. BHPS was funded by the Economic and Social Research Council. Data are available free of charge for non-commercial use from the Economic and Social Data Services. A comprehensive data set documentation is provided by the UK Data Service (Taylor et al., 2018). The question on general life satisfaction appears in waves 6-10 and 12-18 (i.e. 1997-2001 and 2002-2009), therefore we can compute first-difference variables for waves 7-10 and 13-18. We remove 787 individuals who display no within-variation along the panel and report satisfaction at the boundaries (they persistently report $LS = 1$ or $LS = 7$). We end up with 111,894 individual-year observations from 20,269 respondents (53% female, average age 48).

Cross-section linear regression. To single-out the vector of residual life satisfaction which is orthogonal to the observable characteristics of the respondents, we study a linear satisfaction equation and estimate it by OLS. The equation takes the following form: $LS_i = X_i\alpha + u_i$; where LS_i is the life satisfaction of individual i in wave 7, u_i is a normally distributed error term and X_i is a matrix of control variables.

Panel non-linear regression. We study a model where the probability to report a given outcome j, $j = \{$more satisfied, about the same, less satisfied$\}$, is determined by current life satisfaction and other covariates, i.e. we estimate $P(\widetilde{\Delta LS_{it}} = j \mid LS_{it}; X_{it})$, where $\widetilde{\Delta LS_{it}}$ is the recalled change in life satisfaction reported by individual i at time t. We estimate the model by Multinomial Logit and cluster errors at the individual level to correct for individual-specific error correlation over time. Estimation is valid under the IIA assumption, which seems very reasonable in our case. We should also underline that our model contains only case-specific regressors, i.e. variables that do not change according to the decision which is made (reporting a pos., neg. or null change). Whether we had alternative-specific variables (let's say one is paid differently according to what she reports), conditional or mixed logit models would be more appropriate. As a robustness check, we also test a more flexible specification (with dummies for each level of satisfaction, which allows to single out effects at the upper- and lower-bound of the satisfaction scale) and less flexible specification (where we assume the set j=$\{$more satisfied, about the same, less satisfied$\}$ to be ordered and estimate by Ordered Probit). Results, displayed in the SI Appendix (Tables S3 and S4), are substantially the same.

Study 3. *Sample and procedures.* The *Enquête de conjoncture auprès des ménages* (CAMME) is a French monthly consumer confidence survey run by the national statistical office (INSEE). Since June 2016, the survey includes a well-being module, which surveys a representative sample of the French population. Our dataset is a longitudinal series of repeated cross-sections over 11 quarterly waves, from June 2016 to December 2018. The survey is run by telephone, on a representative sample of French residents (overseas territories

excluded), drawn from the housing tax register. The survey is a partnership between INSEE and the Well-being Observatory of CEPREMAP. Data are available free of charge through the Réseau Quetelet. Methodological information is available here: <https://www.insee.fr/fr/metadonnees/source/serie/s1208/presentation>. Our sample consists of 18,589 respondents (54% female, average age 58).

Study 4. *Sample and procedures.* The *Gallup Poll Social Series* is an American survey run by the internationally renowned organization Gallup Inc. Although micro-data are not publicly available, aggregate data were published by the organization both in their complete compilation of polls (Gallup and Newport, 2009) and in a report of the Gallup News Service (Moore, 2006). Our analysis is based on these aggregate figures. Data for table 1 are based on waves 1971, 1972, 1974, 1976, 2001, 2002, 2003, 2004, 2005, 2006. Data for figure 4a are based on waves 1971, 1976, 2001 and 2006 only. In-depth methodological information is available from the Gallup collection (Gallup and Newport, 2009).

Data Availability. All data analyzed in this paper are publicly available and free of charge.

References

Atkinson, T. (1982). The stability and validity of quality of life measures, *Social Indicators Research* **10**(2): 113–132.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*, Macmillan.

Bellet, C., De Neve, J.-E. and Ward, G. (2019). Does employee happiness have an impact on productivity?, *Saïd Business School WP, 13* .

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation, *The Quarterly Journal of Economics* **117**(3): 871–915.

Blanchflower, D. G. and Oswald, A. J. (2008). Is well-being U-shaped over the life cycle?, *Social Science & Medicine* **66**(8): 1733–1749.

Blaney, P. H. (1986). Affect and memory: a review., *Psychological bulletin* **99**(2): 229.

Busseri, M. A., Choma, B. L. and Sadava, S. W. (2009). "as good as it gets" or "the best is yet to come"? how optimists and pessimists view their past, present, and anticipated future life satisfaction, *Personality and individual differences* **47**(4): 352–356.

Busseri, M. A. and Samani, M. N. (2019). Lay theories for life satisfaction and the belief that life gets better and better, *Journal of Happiness Studies* **20**(5): 1647–1672.

Cantril, H. (1965). *Pattern of human concerns*, Rutgers University Press.

Caplin, A. and Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings, *The Quarterly Journal of Economics* **116**(1): 55–79.

Chew, S. H., Huang, W. and Zhao, X. (2019). Motivated false memory, *Available at SSRN 2127795* .

Clark, A. E. (1999). Are wages habit-forming? Evidence from micro data, *Journal of Economic Behavior & Organization* **39**(2): 179–200.

Clark, A. E. (2018). Four decades of the economics of happiness: Where next?, *Review of Income and Wealth* **64**(2): 245–269.

De Neve, J.-E. and Oswald, A. J. (2012). Estimating the influence of life satisfaction and positive affect on later income using sibling fixed effects, *Proceedings of the National Academy of Sciences* **109**(49): 19953–19958.

Easterlin, R. A. (2001). Income and happiness: Towards a unified theory, *The Economic Journal* **111**(473): 465–484.

Fiedler, K. and Hütter, M. (2013). Memory and emotion, *Sage Handbook of Applied Memory* pp. 145–161.

Foster, G., Frijters, P. and Johnston, D. W. (2012). The triumph of hope over disappointment: A note on the utility value of good health expectations, *Journal of Economic Psychology* **33**(1): 206–214.

Frank, R. H. and Hutchens, R. M. (1993). Wages, seniority, and the demand for rising consumption profiles, *Journal of Economic Behavior & Organization* **21**(3): 251–276.

Fredrickson, B. L. and Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes., *Journal of Personality and Social Psychology* **65**(1): 45.

Furnham, A. and Petrides, K. (2003). Trait emotional intelligence and happiness, *Social Behavior and Personality: an International Journal* **31**(8): 815–823.

Gallup, A. and Newport, F. (2009). *Gallup poll: public opinion 2008*, Rowman & Littlefield.

Ganzach, Y. and Yaor, E. (2019). The retrospective evaluation of positive and negative affect, *Personality and Social Psychology Bulletin* **45**(1): 93–104.

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C. and Schupp, J. (2019). The German Socio-Economic Panel (SOEP), *Jahrbücher für Nationalökonomie und Statistik* **239**(2): 345–360.

Goel, V., Rosella, L. C., Fu, L. and Alberga, A. (2018). The relationship between life satisfaction and healthcare utilization: a longitudinal study, *American Journal of Preventive Medicine* **55**(2): 142–150.

Johnson, E. J. and Tversky, A. (1983). Affect, generalization, and the perception of risk, *Journal of Personality and Social Psychology* **45**(1): 20.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A. and Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end, *Psychological Science* **4**(6): 401–405.

Kahneman, D. and Krueger, A. B. (2006). Developments in the measurement of subjective well-being, *Journal of Economic Perspectives* **20**(1): 3–24.

Kaiser, C. (2020). Memories as anchors: Novel analyses on the intrapersonal comparability of wellbeing reports, *SocArXiv* .

Kaplan, R. L., Levine, L. J., Lench, H. C. and Safer, M. A. (2016). Forgetting feelings: Opposite biases in reports of the intensity of past emotion and mood., *Emotion* **16**(3): 309.

Kemp, S., Burt, C. D. and Furneaux, L. (2008). A test of the peak-end rule with extended autobiographical events, *Memory & Cognition* **36**(1): 132–138.

Köke, S. and Perino, G. (2017). For "better" or "worse": a direct approach to elicit preference rankings from life-satisfaction data, *WiSo-HH Working Paper Series* (43).

Lachman, M. E., Röcke, C., Rosnick, C. and Ryff, C. D. (2008). Realism and illusion in Americans' temporal views of their life satisfaction: Age differences in reconstructing the past and anticipating the future, *Psychological Science* **19**(9): 889–897.

Levine, L. J., Murphy, G., Lench, H. C., Greene, C. M., Loftus, E. F., Tinti, C., Schmidt, S., Muzzulini, B., Grady, R. H., Stark, S. M. et al. (2021). Remembering facts versus feelings in the wake of political events, *Cognition and Emotion* pp. 1–20.

Levine, L. J., Prohaska, V., Burgess, S. L., Rice, J. A. and Laulhere, T. M. (2001). Remembering past emotions: The role of current appraisals, *Cognition & Emotion* **15**(4): 393–417.

Loewenstein, G. (2006). The pleasures and pains of information, *Science* **312**(5774): 704–706.

Loewenstein, G. and Sicherman, N. (1991). Do workers prefer increasing wage profiles?, *Journal of Labor Economics* **9**(1): 67–84.

Lyubomirsky, S. (2011). Hedonic adaptation to positive and negative experiences., *in* S. Folkman (ed.), *The Oxford Handbook of Stress, Health, and Coping*, Oxford University Press.

Michalos, A. C. and Kahlke, P. M. (2010). Stability and sensitivity in perceived quality of life measures: Some panel results, *Social Indicators Research* **98**(3): 403–434.

Moore, D. W. (2006). Public optimistic about personal lives in next five years. http://news.gallup.com/poll/22300/public-optimistic-about-personal-lives-next-five-years.aspx.

Oswald, A. J., Proto, E. and Sgroi, D. (2015). Happiness and productivity, *Journal of Labor Economics* **33**(4): 789–822.

Robinson, M. D. and Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report., *Psychological bulletin* **128**(6): 934.

Rosella, L. C., Fu, L., Buajitti, E. and Goel, V. (2019). Death and chronic disease risk associated with poor life satisfaction: a population-based cohort study, *American Journal of Epidemiology* **188**(2): 323–331.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories., *Psychological review* **96**(2): 341.

Saucet, C. and Villeval, M. C. (2019). Motivated memory in dictator games, *Games and Economic Behavior* **117**: 250–275.

Schacter, D. L., Chiao, J. Y. and Mitchell, J. P. (2003). The seven sins of memory: implications for self, *Annals of the New York Academy of Sciences* **1001**: 226–239.

Senik, C. (2008). Is man doomed to progress?, *Journal of Economic Behavior & Organization* **68**(1): 140–152.

Smith, D., Loewenstein, G., Jepson, C., Jankovich, A., Feldman, H. and 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.

Strijbosch, W., Mitas, O., van Gisbergen, M., Doicaru, M., Gelissen, J. and Bastiaansen, M. (2019). From experience to memory: on the robustness of the peak-and-end-rule for complex, heterogeneous experiences, *Frontiers in psychology* **10**.

Taylor, M. F., Brice, J., Buck, N. and Prentice-Lane, E. (2018). *British Household Panel Survey: Waves 1-18, 1991-2009. 8th Edition.*, UK Data Service.

Tversky, A. and Griffin, D. (1991). Endowment and contrast in judgments of well-being', *in* F. Strack, M. Argyle and N. Schwarz (eds), *Subjective Well-Being: An Interdisciplinary Perspective*, Oxford: Pergamon Press.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science* **185**(4157): 1124–1131.

Wilson, A. E. and Ross, M. (2000). The frequency of temporal-self and social comparisons in people's personal appraisals., *Journal of Personality and Social Psychology* **78**(5): 928.

Wilson, A. E. and Ross, M. (2001). From chump to champ: people's appraisals of their earlier and present selves., *Journal of Personality and Social Psychology* **80**(4): 572.

Wilson, T. D., Meyers, J. and Gilbert, D. T. (2003). "How happy was I, anyway?" A retrospective impact bias, *Social Cognition* **21**(6): 421–446.

Zimmermann, F. (2020). The dynamics of motivated beliefs, *American Economic Review* **110**(2): 337–61.