NBER WORKING PAPER SERIES

DO PEOPLE REPORT HAPPINESS ACCURATELY?
A METHOD FOR VALIDATING SURVEY RESPONSES
WITHOUT GROUND-TRUTH BENCHMARKS

James Andreoni
B. Douglas Bernheim
Tingyan Jia

Working Paper 32208
http://www.nber.org/papers/w32208

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2024, Revised July 2026

We are grateful to Sumhith Aradhyula, Renee Qin, and Aniaba N'guessan for able research assistance. We would also like to thank Daniel Benjamin, David Comerford, Leonard Goff, Ori Heffetz, Miles Kimball, Christian Krekel, Nick Netzer, and participants at the LSE Well-Being Seminar for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Do People Report Happiness Accurately? A Method for Validating Survey Responses Without Ground-Truth Benchmarks
James Andreoni, B. Douglas Bernheim, and Tingyan Jia
NBER Working Paper No. 32208
March 2024, Revised July 2026
JEL No. D60, D63

ABSTRACT

We introduce a method for validating survey responses in contexts where ground-truth benchmarks are unavailable. The method involves the measurement of "self-reported misreporting" (SRM). We clarify the inferences one can draw about actual misreporting from SRM under various combinations of assumptions. Then we document the accuracy of SRM through an experiment involving an established paradigm for measuring dishonesty. Applying the method to happiness data, we show that SRM is substantial and varies systematically across experiences. We also confirm that those who expect to exaggerate happiness report greater happiness contemporaneously. We conclude that misreporting of subjective well-being is significant and systematic.

James Andreoni
University of California, San Diego
Department of Economics
and NBER
andreoni@ucsd.edu

Tingyan Jia
University of Leicester
School of Business
Department of Economics
jiatingyan2011@gmail.com

B. Douglas Bernheim
Stanford University
Department of Economics
and NBER
bernheim@stanford.edu

1. Introduction

The recent growth of economic research on self-reported well-being reflects an emerging interest in exploring alternatives to traditional measures of economic well-being. This approach has long been popular in the field of psychology and has gained considerable acceptance within the public policy community. Perhaps the most visible application has involved the construction and refinement of "national happiness accounts" (see, for example, Helliwell et al., 2014; Kahneman et al., 2004).

This paper asks whether people report happiness accurately.[1] Veenhoven (1984) articulates two reasons to fear they may not.[2] First, people "may in their heart know that they are disappointed with life, but repress that thought because they cannot deal with its consequences" (pp. 44-45). Second, they may "tend to be dishonest in their communications on the matter" due to "social desirability bias" (p. 48). This bias could have a variety of sources. For example, people may project happiness to conform with social norms and promote positive interactions. Relatedly, they may fear that others will interpret expressions of unhappiness negatively, as cries for help or admissions of weakness.

Validation of happiness measures is inherently challenging. Various studies have attempted to explore the accuracy of self-reported well-being but the existing evidence is inherently limited. The central difficulty is that subjective sensations are unobserved. Critically, a good solution to this problem would have potential value in a wide range of applications where definitive validation is impossible because "ground truth" benchmarks are unavailable. Examples include survey responses to questions concerning (i) willingness to pay for hypothetical environmental goods (contingent valuation); (ii) time allocation, particularly for "virtuous" activities such as exercise, news consumption, and household budgeting; (iii) the consumption of "sin goods" such as cigarettes, alcohol, and recreational drugs; (iv) perceptions of social norms; (v) workplace satisfaction; (vi) the frequency of dishonest behaviors such as cheating and theft; (vii) aspects of personal relationships such as satisfaction, fidelity, or frequency of conflict; and (viii) political opinions.

This paper develops and implements a novel validation strategy involving the measurement of "self-reported misreporting" (SRM). The approach works as follows. First, we direct people to envision a hypothetical scenario in which the quantity of interest would be q. Then we ask them to indicate the response they think they would give if asked to report the quantity on a survey under precisely the same

1 Typical survey questions reference a variety of related concepts, such as happiness, life satisfaction, and domain-specific satisfaction. Our analysis is not specific to any particular concept. Throughout this paper, we will use the term "happiness" as shorthand for any notion of self-assessed well-being.

2 To be clear, these concerns relate to *misreporting*. Veenhoven (1984) mentions other reliability concerns including the possibilities that people may have no actual opinion, that they may not know their true happiness, that their responses may be inconsistent and unstable, that they may give stereotyped responses, that interviewers may impart biases, and that responses may be sensitive to context and methods. See also Veenhoven (2008, 2015).

conditions. SRM is the difference between q and the value they say they would anticipate reporting. One can either fix q as part of the scenario's description or, to ensure that respondents find the hypothetical plausible, ask them to indicate the level of q they think would prevail under the specified conditions before asking them what they would report.

More concretely, to evaluate the accuracy of data on happiness, we begin by selecting a set of reasonably common events, both positive and negative. We instruct respondents to imagine that one of the events occurs at some specified point in the future, and pose two questions: First, how happy do they think they would feel? Second, assuming they were asked in a survey how happy they felt under precisely the same conditions, how do they think they would respond? The discrepancy between the two answers measures SRM for happiness.

Our first objective is to clarify precisely what one learns from SRM measures. Absent any further assumptions, non-zero SRM implies that the respondent either expects to misreport q or misrepresents the value they expect to report (or both). Accordingly, to credit their statements about the quantity of interest, one must embrace one of the following two assumptions: either (i) the respondent is truthful about q even though they believe they are not, or (ii) the respondent happens to be honest when reporting q even though they are dishonest when making related statements. Because these assumptions are unappealing, non-zero SRM undermines confidence in reported values of q. Stronger conclusions follow under various combinations of additional assumptions. For example, we identify assumptions under which one can treat SRM as a lower bound on the respondent's actual expectation concerning misreporting, as an estimate of that expectation, as a lower bound on actual misreporting, and as an estimate of actual misreporting. We also explain why it may be reasonable to think these various assumptions are satisfied.

It is possible to test these assumptions and thereby validate our validation method in settings where ground truth is known for the population but not for individuals. To this end, we conduct an experiment based on an established paradigm for measuring the extent to which people lie when only they know the truth (Fischbacher and Föllmi-Heusi, 2013). Participants privately observe the realization of a random variable and are asked to report it, knowing that the reported value determines their bonus. Using this framework, we measure misreporting for a sample of participants by comparing the average report to the known mean of the random variable. For a second sample, we ask participants to say what they think they would report conditional on specified realizations. We use those data to construct SRM. We find that average SRM for the second sample is nearly identical to average misreporting for the first sample. While these findings are specific to the context studied, they lend credibility to the assumptions that justify using SRM as a measure of misreporting rather than merely as a bound or diagnostic.

The rest of this study uses our method to evaluate the accuracy of self-reported happiness. We seek to determine not only whether people misreport happiness, but also whether misreporting varies systematically according to the types of conditions the individual experiences. Such variation would raise questions concerning the usefulness of self-reported well-being in contexts where the analyst seeks to compare welfare under distinct conditions. We document substantial self-reported misreporting. While SRM tends to be larger for negative events than for positive events, there is no simple pattern such as consistent overstatement, understatement, or compression. The variation in SRM implies that comparisons of happiness across conditions are problematic.

Even though we cannot fully validate our method within the context of happiness given the absence reliable population benchmarks, we can partially validate it by exploring whether the mechanisms that drive expected misreporting of conditional happiness also infect reports of current happiness. Specifically, we show that people who expect to exaggerate their happiness when asked about a future contingency also report greater happiness when asked about their current state of mind.

The questions we pose concerning expected happiness contingent on future events resemble those used in studies that construct measures of projection bias (see, e.g., Gilbert et al., 1998, and Loewenstein et al., 2003). Accordingly, as a further application, we use our results to estimate the extent to which measured projection bias reflects misreporting of happiness once events actually occur. While evidence of projection bias remains, a substantial portion of the effect appears to be spurious.

The rest of the paper proceeds as follows. Section 2 describes related literature. Section 3 elaborates on our method, clarifies the conclusions one can draw from SRM under various combinations of assumptions, explains why those assumptions may be plausible, and tests the approach in a setting where validation is feasible. Section 4 presents our analysis of self-reported happiness. It describes our surveys, presents our basic findings on self-reported misreporting, provides evidence that the mechanisms driving misreporting infect reports of current happiness, and explains how our results impact the interpretation of evidence on projection bias. Section 5 elaborates on some other potential applications of the method and offers some brief conclusions.

2. Related Literature

This study is most closely related to three strands of literature on the measurement of subjective well-being. The first seeks to evaluate the accuracy of self-reported happiness; the second explores the limitations of welfare measures based on discrete Likert scales; the third asks whether people attempt to maximize self-reported happiness. We review each in turn.

Research in hedonic psychology has provided most of the existing evidence concerning the accuracy of self-reported happiness. Validation efforts have involved two main strategies. One is to investigate whether such reports align with other measures of happiness, such as non-verbal cues (e.g., amount of smiling), expert ratings by psychologists, interviewer ratings, reports of family and friends, recollection of positive events, and various biometric responses. Correlations among these measures are said to provide "congruent validation." The second strategy is to examine correlations with other phenomena that are generally thought to connect with happiness, such as measures of social success and other aspects of "the good life." These correlations are said to provide "concurrent validation." For reviews of this evidence, see Veenhoven (1984, 2008, 2015) and Diener and Oishi (2004); with respect to biometric evidence, see Larsen and Frederickson (1999) and Davidson (2004).

A fundamental limitation of these validation methods is that they attempt to evaluate a happiness proxy by comparing it with other proxies rather than with ground truth (i.e., "true happiness").[3] This limitation is intrinsic because, in this context, ground truth is never observed. Accordingly, one cannot rule out the possibility that correlations across indicators of happiness reflect interrelated biases rather than variations in true happiness. Moreover, even if we construe this evidence as establishing the existence of a correlation between reported and true happiness, it cannot establish the absence of systematic bias. Formally, if reported happiness under condition x is given by $r(x) = h(x) + b$ where $h(x)$ is true happiness and b is the reporting bias, then even a strong positive correlation between $h(x)$ and $r(x)$ does not imply that b is white noise. On the contrary, b may vary systematically with x, in which case reports may reverse rankings for categories of events.

Other evidence attempts to speak directly to the hypothesis that people overstate happiness. First, reported happiness is "slightly higher in personal interviews" than on questionnaires, for which social pressures are perhaps less salient (Veenhoven, 2008). While suggestive, that comparison does not control for differences between the samples—e.g., for the possibility that those who consent to personal interviews are a bit more social and consequently happier. Nor does it establish whether bias remains in questionnaires.

Second, some studies have asked whether reported happiness correlates with personal characteristics that are plausibly associated with misreporting, such as general defensiveness or a desire for social approval. For a survey of these results, see Veenhoven (1984), who describes them as mixed and concludes that such correlations are modest (pp. 46, 50). Because true happiness is also presumably related to personality traits, the implications of these findings are far from clear. For example, some studies construe helpfulness as indicating a need for social approval. But helpful people may report greater happiness because happy people tend to be helpful, or because helping others contributes to happiness.

Third, on average, people characterize their level of happiness as above average (e.g., Goldings, 1954). While this finding is consistent with the hypothesis that survey responses tend to overstate happiness, there are other possible explanations. As noted in the literature on overconfidence, such findings are

[3] Veenhoven (1984, p. 53) makes the point as follows: "This approach obviously requires that we must be sure about the validity of at least one happiness indicator: the measure that is used as a standard. Unfortunately this requirement cannot be met. We are not sure of the validity of any."

potentially consistent with Bayesianism. One possibility is that different people assess happiness differently, placing the greatest weight on the dimensions of experience from which they personally derive the greatest happiness (analogously to the mechanism described in Santos Pinto and Sobel, 2005). Another possibility is that the pattern reflects asymmetries in the underlying signal structure (analogously to the mechanism described in Benoit and Dubra, 2011). Stepping outside the Bayesian paradigm, it is also possible that people underestimate the happiness others feel, possibly because popular media coverage skews toward misfortune.

Considering the limitations of these efforts to validate self-reported happiness, the hedonic psychology literature has been surprisingly sanguine concerning its use. For example, after acknowledging that "[m]any misgivings have been advanced about such self-report of happiness," Veenhoven (2015, p. 382) concludes that "[t]hese qualms have not been supported by empirical research…" (p. 383; similar statements appear in Veenhoven 1984, 2008). Likewise, with respect to evaluating the quality of life as a whole, Van Praag and Ferrer-i-Carbonell (2004) observe that respondents "apparently have no difficulty in answering such a question, and… those responses seem to be comparable. Hence, we will accept this as empirical evidence that respondents are able to evaluate their life and that those responses lend themselves to scientific analysis."

The second strand of literature focuses on measurement issues other than misreporting. One concern is that the use of categorical Likert scales invalidates comparisons of average happiness across groups (Bond and Lang, 2019). Because the cardinalization of happiness is arbitrary, one can legitimately say that average happiness is higher for one group than for another only by establishing first-order stochastic dominance with respect to the underling happiness states. Unfortunately, when observations are categorical, and when the highest and lowest response categories are non-empty, one cannot draw non-parametric conclusions about stochastic dominance. This observation calls the usefulness of standard happiness data into question. Liu and Netzer (2023) propose using response times to remedy this problem.

The third strand of literature explores the relationship between anticipated happiness and choice. It finds that, when choosing, people do not try to maximize self-reported happiness (Benjamin et al., 2012, 2014; Bernheim, Kim, & Taubinsky, forthcoming). A common interpretation of this finding is that happiness is not the sole object of desire and consequently not an appropriate gauge of economic well-being. There are three potential objections to this interpretation. First, the evidence does not rule out the possibility that desire is limited to happiness, which self-reported happiness systematically mismeasures.[4] The current

[4] Bernheim, Kim, and Taubinsky (forthcoming) confirm that choice loads on anticipated sensations other than happiness even when one accounts for unsystematic measurement error. However, they do not exclude the possibility that this finding reflects systematic error in measuring happiness.

study addresses that issue directly. Second, a fundamental premise of behavioral economics is that choices do not reliably reflect desires. Accordingly, one can reinterpret evidence of conflicts between choice and broad expressions of subjective well-being as indicting the former rather than the latter (see, e.g., Stutzer & Frey, 2008). Third, and more fundamentally, conclusions about welfare depend on one's philosophical school of thought. For the most part, economists implicitly assume that "well-being consists in having one's preferences satisfied" (Kagan, 1998, p. 36). But this perspective, known as Desire Satisfaction Theory, is philosophically controversial (see, e.g., Parfit, 1984, Hausman, 2012). Those who favor the use of self-reported happiness as a measure of well-being often subscribe, either implicitly or explicitly, to an alternative (and equally controversial) school of thought known as Welfare Hedonism. The guiding principle of that perspective is that "well-being consists solely in the presence of pleasure and the absence of pain" (Kagan, 1998, p. 31). Within economics, this view is commonly associated with Bentham (2023) and Mill (2012). Critically, Welfare Hedonists deny the importance of desire except insofar as its satisfaction leads to happiness, broadly construed. For them, the accurate measurement of happiness is critical even if it is not a sufficient statistic for choice.

3. Methods

In this section, we introduce the concept of self-reported misreporting and explain its interpretation under various combinations of assumptions (Section 3.1). Then we present evidence addressing the validity of those interpretations (Section 3.2).

3.1. Conceptual framework

Suppose we are interested in assessing the accuracy with which people report some hard-to-validate quantity q that varies with conditions w according to the function $q = Q(w)$. Examples include the intensity of a subjective sensation such as happiness, willingness-to-pay for an environmental good, cash tips earned by service staff at restaurants, and time spent on household budgeting. When asked about this quantity once conditions w have arisen, the individual reports $r = R(q, w)$. We are concerned with detecting the possibility that $R(q, w) \neq Q(w)$ and, if possible, gauging the degree of misreporting. While we may observe r conditional on various events, we do not observe the corresponding q even in the aggregate, which means we cannot evaluate the accuracy of r directly.

There are two versions of our method, a basic version and an extended version. Both versions involve the elicitation of responses concerning hypothetical rather than realized conditions. Because complete hypotheticals are difficult to describe, we divide w into two components: conditions the hypotheticals specify (x), and conditions they leave unspecified (z). In some applications, z might include both external circumstances and various internal emotional reactions.

The basic method. For the basic method, we pose hypothetical scenarios consisting of a partially specified event, x^h, and an associated quantity, q^h. Implicitly, specifying both x^h and q^h calls to mind some additional conditions $z^h = Z(x^h, q^h)$ that allow the individual to make sense of the hypothetical. Formally, $Z(x^h, q^h)$ satisfies $q^h = Q^E(x^h, Z(x^h, q^h))$, where $Q^E(w)$ represents the respondent's expectation of q given the complete hypothetical scenario w.[5] Significantly, we do not assume the respondent necessarily understands the process that generates q (in other words, that Q and Q^E coincide). When there exists no reasonable z^h for which $q^h = Q^E(x^h, z^h)$, the respondent finds the partially specified hypothetical (x^h, q^h) implausible. The motivation for our extended method, discussed below, is to ensure that the analyst does not inadvertently pose an implausible (incomplete) hypothetical.

We then ask the individual to indicate the quantity q they would report in the hypothetical scenario. Let $r^E = R^E(x^h, q^h)$ denote the quantity they truly think they would report and let $r^e = R^e(x^h, q^h)$ denote what they say they think they would report. To be clear, both $R^e(x^h, q^h)$ and $R^E(x^h, q^h)$ account for the additional conditions $Z(x^h, q^h)$ that come to mind when considering the partially specified hypothetical. *Self-reported misreporting* for this hypothetical is then the difference between the specified quantity and the quantity they say they think they would report:

$$SRM(x^h, q^h) = q^h - R^e(x^h, q^h).$$

We define *actual misreporting* as the difference between the specified quantity and the quantity the individual would report if the specified scenario actually materialized with q^h unobserved:

$$MR(x^h, q^h) = q^h - R(q^h, x^h, Z(x^h, q^h))$$

$MR(x^h, q^h)$ is the object we would ideally like to measure but cannot measure because, in any realized scenario, the quantity of interest is unobservable. When evaluating the credibility of responses concerning that quantity, it may also be relevant to know whether people truly believe they will misreport it. We therefore define *expected misreporting* as the difference between the specified quantity and the quantity the individual truly thinks they would report:

$$EMR(x^h, q^h) = q^h - R^E(x^h, q^h),$$

What inference can we draw, if any, concerning expected misreporting, $EMR(x^h, q^h)$, and actual misreporting, $MR(x^h, q^h)$, from the observation that self-reported misreporting is non-zero, $SRM(x^h, q^h) \neq 0$? Even without additional structure or assumptions, we can infer that either

[5] Uniqueness of the solution is inessential; the idea is simply that there is some way for the individual to complete the hypothetical.

$EMR(x^h, q^h) \neq 0$, $R^e(x^h, q^h) \neq R^E(x^h, q^h)$, or both. The first possibility implies that the individual believes their answers to questions about q are not reliable. To credit those answers, one must therefore disbelieve the respondent's assessment of their own reliability. The second possibility impugns the individual's honesty within the domain of interest. To credit their answers about q, one must therefore assume they are selectively and fortuitously honest. In other words, with either possibility, one requires an unappealing assumption to justify taking statements about q at face value.

We can reach stronger conclusions under various combinations of the three assumptions listed below. The first is that $R^e(x^h, q^h)$ lies between $R^E(x^h, q^h)$ and q^h:

Assumption 1: $R^e(x^h, q^h) \in [\min\{R^E(x^h, q^h), q^h\}, \max\{R^E(x^h, q^h), q^h\}]$

The logic of this assumption is that the individual is likely averse both to lying and to admitting that they lie. To avoid lying, they would report $r^e = R^E(x^h, q^h)$. To avoid admitting that they lie, they would report $r^e = q^h$. Provided the costs of lying and admitting to lying are both increasing in the size of the lie, the optimal report falls between these two boundary cases.[6]

Assumption 1 immediately implies that $R^e(x^h, q^h) - q^h$ and $R^E(x^h, q^h) - q^h$ have the same sign and that the second expression is larger in absolute value than the first. Consequently, when $SRM(x^h, q^h) \neq 0$, we have:

$$\frac{EMR(x^h, q^h)}{SRM(x^h, q^h)} \geq 1$$

In other words, under Assumption 1, self-reported misreporting (weakly) understates the degree to which the respondent thinks they would misreport q.

The next assumption is that people correctly anticipate what they would say in the completed hypothetical scenario:

Assumption 2: $R^E(x^h, q^h) = R(q^h, x^h, Z(x^h, q^h))$.

The logic of this assumption is either that people can determine what they would say through introspection or that they know their own proclivities from experience. To be clear, the assumption does *not* require them to correctly anticipate the unspecified conditions, z, that would likely accompany the

[6] Formally, suppose the individual's utility is given by $U(f(r^e), g(r^e))$, where U is strictly increasing in both its arguments, f is single-peaked with a maximum at q^h (i.e., lower values correspond to admitting a greater lie), and g is single-peaked with a maximum at r^E (i.e., lower values correspond to telling a greater lie). Consider the case where $q^h > r^E$. Then, for any $r > q^h$, we have $f(r) < f(q^h)$ and $g(r) < g(q^h)$, which together imply $U(f(r), g(r)) < U(f(q^h), g(q^h))$. A similar argument establishes the suboptimality of $r < r^E$. The case of $q^h < r^E$ is symmetric.

specified conditions, (x^h, q^h), in practice. Accordingly, in applications where z includes unspecified emotional states, there is no requirement that people can accurately forecast how they will feel.[7]

When we observe $SRM(x^h, q^h) \neq 0$, Assumption 2 implies that either $R(x^h, Z(x^h, q^h)) \neq q^h$ (equivalently, $MR(x^h, q^h) \neq 0$), $R^e(x^h, q^h) \neq R^E(x^h, q^h)$, or both. In other words, misreporting is definitely present within the domain of interest: either the individual would misreport q^h if asked to do so upon the realization of conditions $(x^h, Z(x^h, q^h))$, or they misreport their expectation of what they would say if asked to disclose q^h under the same conditions when posed hypothetically. Although we cannot pin down where the misreporting lies, its existence casts doubt on the respondent's honesty within this domain, and hence on the reliability of $R(q, x, Z(x, q))$ as a measure of q.

If Assumptions 1 and 2 are both satisfied, then we can conclude that $R^e(x^h, q^h) - q^h$ and $R(q^h, x^h, Z(x^h, q^h)) - q^h$ have the same sign and that the second expression is (weakly) larger in absolute value than the first. It follows that

$$\frac{MR(x^h, q^h)}{SRM(x^h, q^h)} \geq 1$$

In other words, self-reported misreporting in response to a hypothetical concerning q (weakly) understates the degree of misreporting when the completed hypothetical scenario actually materializes and q is unobserved.

Our third assumption strengthens the first by elevating a boundary case:

Assumption 3: $R^e(x^h, q^h) = R^E(x^h, q^h)$.

The logic of this assumption builds on the intuition for Assumption 1. The desire to avoid lying draws the individual toward $r^e = R^E(x^h, q^h)$, while the desire to avoid admitting that they would expect to lie under hypothetical circumstances draws them toward $r^e = q^h$. Arguably, admitting an actual lie is much worse than acknowledging the hypothetical possibility of telling a lie. In that case, the first effect should dominate.

Assumption 3 would also hold under the more restrictive hypothesis that the psychological mechanisms responsible for reporting biases do not apply to hypotheticals. For example, in the case of

7 To illustrate, suppose we want to evaluate whether people exaggerate the amount of time they spend reading about gubernatorial candidates prior to elections. We might describe a hypothetical scenario in which the individual spends 45 minutes reading about two candidates, providing a short bio for each. The respondent might complete the hypothetical by imaging a combination of external circumstances and internal emotional reactions they consider plausible that rationalize the 45-minute commitment. Assumption 2 does not require that those circumstances and emotional reactions are likely to arise in an actual election. For instance, it is consistent with the possibility that people systematically overestimate how much they will care about elections.

happiness, Veenhoven (1984) articulates two reasons for misreporting: repression of feelings of disappointment with life and social influences. The psychological motives for repression, such as ego defense, plainly apply to contemporaneous feelings but not necessarily to anticipated feelings associated with an event that may not occur. Likewise, people may be less likely to interpret statements about contingent future feelings as cries for help or expressions of weakness. More generally, the norms that govern social interactions call for the appearance of contemporaneous happiness, rather than happiness at points in time and under conditions far removed from the interaction.

Under Assumption 3, self-reported misreporting accurately measures the degree to which the person thinks they would misreport: $SRM(x^h, q^h) = EMR(x^h, q^h)$. Moreover, the combination of Assumptions 2 and 3 implies that self-reported misreporting accurately measures misreporting: $SRM(x^h, q^h) = MR(x^h, q^h)$.

The following proposition, though mathematically trivial, usefully summarizes these observations:

Proposition: *Suppose* $SRM(x^h, q^h) \neq 0$. *Then either the individual believes their answers to questions about* q *are unreliable,* $EMR(x^h, q^h) \neq 0$, *or they misreport the responses they anticipate giving,* $R^e(x^h, q^h) \neq R^E(x^h, q^h)$ *(or both). In addition:*

(i) *Under Assumption 1, self-reported misreporting provides a lower bound on expected misreporting:* $\frac{EMR(x^h,q^h)}{SRM(x^h,q^h)} \geq 1$.

(ii) *Under Assumption 2, the individual either misreports the response they anticipate giving in the specified hypothetical scenario,* $R^e(x^h, q^h) \neq R^E(x^h, q^h)$, *or misreports the quantity of interest when the specified scenario actually materializes and* q *is unobserved,* $MR(x^h, q^h) \neq 0$ *(or both).*

(iii) *Under Assumptions 1 and 2, self-reported misreporting provides a lower bound on actual misreporting:* $\frac{MR(x^h,q^h)}{SRM(x^h,q^h)} \geq 1$.

(iv) *Under Assumption 3, self-reported misreporting accurately measures expected misreporting:* $SRM(x^h, q^h) = EMR(x^h, q^h)$.

(v) *Under Assumptions 2 and 3, self-reported misreporting accurately measures actual misreporting:* $SRM(x^h, q^h) = MR(x^h, q^h)$.

The extended method. The basic version of our method is well-suited for applications in which the analyst knows the range within which q^h might plausibly fall under specified hypothetical conditions, x^h. Because our objective is to validate reports about q in applications where true values are unobserved, the plausible range may be difficult to determine. As a result, the analyst may inadvertently pose a partially specified hypothetical (q^h, x^h) that the respondent considers implausible, in the sense defined earlier (i.e., there is no reasonable z for which $q^h = Q^E(x^h, z)$). In that case, the response to the hypothetical may be unrepresentative of how the individual truly thinks and behaves. As an example, suppose we ask respondents to indicate how happy they would report feeling conditional on some hypothetical negative event (x^h) assuming their honest answer would be 4 on a seven-point Likert scale (q^h).[8] Someone who is highly averse to the event in question may have difficulty understanding how 4 could be an honest answer. From that individual's perspective, the question asks them to indicate what they would report if they experienced the event as someone else might experience it rather than as they themselves can easily imagine experiencing it. This disconnect may distance their answer from their own proclivities.

The extended version of our method addresses such concerns. Instead of specifying a hypothetical scenario consisting of both x^h and q^h, we only specify x^h. Then we ask people to say what they expect the value of q would be if, hypothetically, the event x^h materialized ("the expectations question"). Finally, we ask them to indicate what they would expect to report conditional on the same event, meaning the scenario consisting of x^h and the value of q they have supplied ("the reporting question"). Self-reported misreporting equals the difference between these two responses.

The only distinction between the basic and extended methods is that, in the latter, the respondent first supplies q^h. The advantage of the extended method is that the respondent's answer to the expectations question is more likely to establish a hypothetical (q^h, x^h) they consider plausible for themselves, meaning they can complete it with some reasonable $Z(q^h, x^h)$ for which $q^h = Q^E(x^h, Z(q^h, x^h))$.

To be clear, the extended method does not assume that the respondent's expectations are accurate (i.e., that $Z(q^h, x^h)$ is in fact reasonable, or that Q^E and Q coincide). Nor does it assume that respondents answer the expectations question "truthfully," at least in the usual sense of this term. Even a truthful participant could potentially provide a range of responses because the incomplete hypothetical leaves z unspecified. The method only assumes that the respondent reports a value of q they can imagine experiencing when x^h occurs. That property allows them to envision a completed hypothetical $(q^h, x^h, Z(q^h, x^h))$ they consider internally consistent (in the sense that $q^h = Q^E(x^h, Z(q^h, x^h))$) when answering the reporting question.[9]

---

[8] Because standard survey questions measure happiness on a unitless scale, there is some ambiguity about the meaning of "honest" or "accurate" reporting. We assume each individual adopts a fixed normalization for internally gauging degrees of happiness. When we say that their reporting is accurate, we mean that their quantitative responses are appropriate given this normalization.

[9] Returning to the example from footnote 7, a respondent who wishes to exaggerate the amount of time they would spend reading about the candidates in response to the expectations questions does not need to be "untruthful." Instead, they can implicitly complete the hypothetical by imagining a higher-than-usual level of emotional engagement with the election. The reporting question can then call to mind a level of engagement that rationalizes the expectation they reported.

The order of the questions. The logic of the extended method is clearest when the expectations question comes *first* and the reporting question is *unanticipated*. This protocol offers several advantages. First, it permits the reporting question to reference a fixed quantity q^h, and it enables the analyst to observe that quantity. Second, it removes any incentive for respondents to state an expectation q^h that artificially achieves a high degree of alignment between q^h, $R^E(x^h, q^h)$, and $R^e(x^h, q^h)$. Such alignment might potentially benefit the respondent by minimizing both the need for and appearance of dishonesty when answering the reporting question. Third, the protocol allays potential concerns that respondents might distort their answers to the reporting question to signal their expectations. After answering a direct question about their expectations, respondents are unlikely to view their answer to the reporting question as a further signal of that expectation.

Conceivably, the method might also work if we reversed the order of the questions: the respondent might answer the reporting question by naming a response they could imagine making for some reasonably likely z^h conditional on x^h occurring, and then answer the expectations question by reporting $Q^E(x^h, z^h)$, the value of q they would expect to experience for that same scenario. However, we have less confidence in these assumptions. When the reporting question comes first, respondents may see their answer as providing information about their expectation concerning q. That consideration may pull their answer toward their expectation of q for the z they have imagined. Whether they then answer the subsequent expectations question for that z or some alternative z' that justifies this distorted answer is unclear. Moreover, even if they have the event (x^h, z^h) in mind, they may not report the associated level of q truthfully. For example, once they say what they think they would report, their aversion to acknowledging dishonesty may pull their subsequent answer to the expectations questions toward their prior answer to the reporting question. We therefore conjecture that reversing the order of the questions will artificially reduce the measured magnitude of self-reported misreporting. The design of our main experiment allows us to investigate this conjecture.

Retrospective questions. Past experiences also offer opportunities for assessing self-reported misreporting. A limitation is that one can only ask people about events they have, in fact, experienced. To ensure adequate sample sizes, such events would typically consist of some reasonably common condition x rather than a specific value of q. In the spirit of our extended method, one would ask each respondent who has experienced x to report the associated level of q (the "recollection question") as well as the level they would have reported if asked; self-reported misreporting equals the difference. To the extent recollections are hazy and people flexibly reconstruct significant portions of their memories, they may treat the recollection question much like a hypothetical question about a future event, in which case the method would have the same properties. However, when someone's memory pins down the entire event (q, x, z), the method may work less well, especially if the respondent misreports q as $q' \neq q$ in response to the recollection question. If they then answer the reporting question with respect to the actual event (q, x, z), there will be a mismatch: self-reported misreporting will get at $q' - R^e(x, q)$ rather than $q' - R^e(x, q')$ or $q - R^e(x, q)$. Alternatively, they may try to answer the reporting question with respect to the misrepresented event (q', x, z). However, because that event is infeasible, they may have difficulty imagining q' retrospectively without rewriting history with respect to z. Whether they would reliably envision such counterfactual histories is unclear. Despite these limitations, our analysis of happiness in Section 4 explores the use of both prospective and retrospective data.

3.2. Validating the validation method

Next, we test the assumptions that justify our method in a domain for which validation is feasible. Regrettably, direct validation is not feasible in the context of self-reported happiness because true happiness is fundamentally unobservable. Still, demonstrations of the method's reliability in domains that permit such assessments boost confidence in other applications.

Validation requires us to compare self-reported misreporting to actual misreporting. This requirement is difficult to satisfy because we wish to know how our method performs in settings where respondents know that their true value of q is unobservable. Our strategy is to validate the basic version in a setting where the truth is indeed unknown for individuals, but the distribution of outcomes is known for the group.[10]

For this purpose, we conducted an experiment consisting of two components surveys. The first is, effectively, a replication of previous studies, beginning with Fischbacher and Föllmi-Heusi (2013), that measure the extent to which people lie when only they know the truth: we ask each online participant to flip a coin 10 times, and then to report the number of heads, knowing that they will receive a bonus equal to that number times \$0.10. For example, if they report flipping eight heads, they receive a bonus of \$0.80. While we cannot observe the number of heads for any given participant, we know the true population distribution and can test whether it coincides with the distribution of reports.

The second component survey implements our basic method for assessing reports through self-reported misreporting. The basic method is suitable for this application because the range of plausible outcomes is known. Specifically, we describe the first survey to a second group of participants and ask what they think they would report conditional on various outcomes of the coin flips, which here serve as both x^h and q^h (in other words, there are no unspecified conditions, and Q is the identity function). We then compare the distribution of anticipated reports to the distribution of actual reports, as well as to the true population distribution.

[10] We are grateful to Nick Netzer for suggesting the general strategy of validating the method in a domain for which the truth is known.

The first survey starts with a screen that asks the participant to flip a coin ten times, and to record the number of heads. In the second survey, participants view the same screen and are then asked whether they think they would actually flip a coin.

The next screen in the first survey describes the bonus scheme and asks the participant to report the number of heads. It includes the following statement: *We would appreciate an honest answer*. In the second survey, participants view the same screen. Those who previously said they would not flip a coin are asked whether they would do so at this point in the survey. All participants in both surveys also answer an understanding question concerning the incentive scheme.

Participants in the second survey then see a screen on which they are asked to indicate how they think they would have answered the question concerning the outcome of flipping the coin 10 times. For those who say they would actually flip a coin, we ask:

> *Imagine that you flipped the coin 10 times, and that it has come up Heads [N] times. What do you think you would answer concerning the number of times the coin came up Heads, out of 10, if you were actually a participant in the study? We would appreciate your honest response.*

We supply the number N to each participant, mimicking the population distribution of coin flips. For those who say they would not actually flip a coin, we ask:

> *Given that you would not have flipped the coin 10 times, what do you think you would have answered concerning the number of times the coin came up Heads, out of 10? We would appreciate your honest response.*

We fielded the two surveys on Prolific in November 2023. In total, 196 respondents completed the first survey and 209 completed the second.[11]

The mean of the true population distribution for the number of heads out of 10 flips of a fair coin is 5.0. In contrast, the mean reported number of heads in the first survey is 5.62 ($SE = 0.13$). We strongly reject the hypothesis that people respond truthfully ($p < 0.001$). Instead, we find that the typical respondent exaggerates the number of heads by 0.62. In the second survey, the mean number of heads participants say they would report in such a survey is 5.57 ($SE = 0.13$). We strongly reject the hypothesis that people say they would respond truthfully ($p < 0.001$). Instead, we find that the typical respondent says they would exaggerate the number of heads by 0.57.[12] Remarkably, the difference between misreporting (0.62) and self-reported misreporting (0.57) is only 0.05 ($SE = 0.18$).[13] According to the point estimates, self-reported

---

[11] The full text of each survey appears in the accompanying "Survey Instruments" document. We conducted this study under human subjects protocol IRB-42264 approved by Stanford University's IRB.

[12] The difference between the mean number of heads participants say they would report, and the mean number of simulated heads for those participants, is slightly larger (0.65).

[13] Initially we collected 98 responses to the first survey and 105 to the second. The mean reported number of heads was 5.69 ($SE = 0.18$) in the first survey, and the mean number of heads participants said they would report was 5.63 ($SE = 0.20$) in the second survey. We then doubled the sample size to check that the small difference between these means was not a fluke.

misreporting slightly understates misreporting, but we cannot reject the hypothesis that the two are identical ($p = 0.79$).

Figure 1 shows CDFs for (i) the true population distribution for the number of heads out of 10 flips of a fair coin (in blue), (ii) the distribution of reports from the first survey (in red), and (iii) the distribution of predicted responses from the second survey (in green). The second and third of these CDFs are similar but not identical. However, both are plainly right-shifted relative to the first; the second distribution is weakly higher than the first in the sense of first-order stochastic dominance, and the same is nearly true for the third versus the first.[14]

Overall, 85% of the participants in our second survey said they would actually flip a coin if they participated in the first survey. For those subjects, we can compare the hypothetical outcomes we gave them to their predicted responses. The average simulated number of heads is 4.88, and the average predicted response is 5.44. The average difference for this group, 0.56, is only slightly less than the overall average (0.65).[15] Self-reported misreporting declines sharply as the hypothesized outcome becomes more favorable. Subjects say they would exaggerate by an average of 1.16 flips ($SE = 0.29$) if the hypothesized outcome is less than 4, by an average of 0.92 flips ($SE = 0.17$) if the hypothesized outcome is 4 or 5, and by an average of 0.05 flips ($SE = 0.16$) if the hypothesized outcome is 6 or greater.

These results confirm that $\frac{MR(x^h,q^h)}{SMR(x^h,q^h)} \geq 1$, consistent with the implications of Assumptions 1 and 2 (part (iii) of the Proposition). Indeed, they are roughly consistent with the equality of self-reported misreporting and actual misreporting, consistent with the implications of Assumptions 2 and 3 (part (v) of the Proposition).

4. Application to happiness

Next, we use our method to evaluate the reliability of self-reported happiness. Section 4.1 details our study design. Section 4.2 provides measures of self-reported misreporting for a collection of hypothetical and recollected events. Section 4.3 investigates the relationship between self-reported misreporting and actual reporting of contemporaneous happiness. Section 4.4 explores implications for the phenomenon of projection bias.

---

[14] For zero heads and one head, the cumulative density of predicted responses from the second survey is slightly greater than the cumulative density for coin flips.

[15] The average response for those who said they would not flip a coin is 6.28. As a result, this group is more inclined to exaggerate relative to the simulated distribution (by 1.16, rather than 0.56). Although the group who said they would not flip a coin is small (32 subjects), the difference between the average response for this group and those who said they would flip a coin is nearly statistically significant at the 5% level ($p < 0.056$).

4.1. Study design

We gathered data on happiness in two waves, which we will call our *main study* and our *supplemental study*. The main study gathered the data required to assess self-reported misreporting for four categories of life events, two negative and two positive. Negative events include the end of an existing romantic relationship (for those currently in relationships) and the death of a loved one. Positive events include receiving a promotion at work (for currently employed subjects) and being hired for a dream job (for currently unemployed subjects).

We fielded three short surveys on Prolific between April and May 2023, one for each of the negative events, and one for the two positive events. In total, 1004 respondents participated in the survey on romantic relationships, 300 in the one on loss of a loved one, and 302 in the one on employment.[16]

All three surveys begin with a standard elicitation of current happiness:

> *On a scale from 1 to 7, where 1 is not happy and 7 is very happy, how do you feel now on a typical day?*

Next, we pose questions that separate respondents into pertinent groups. For the survey on romantic relationships, we ask whether the respondent is currently in a relationship. Those who are in relationships answer the prospective breakup questions, while those who are not in relationships answer the retrospective questions if a breakup occurred within the previous six months.[17] For the survey on death of a loved one, we ask whether the respondent experienced such a loss within the previous six months. Those who had that experience answer the retrospective loss questions, while those who did not have that experience answer the prospective loss questions. For the employment survey, we first ask whether the respondent is employed. Those who are not employed and who indicate a potential interest in finding a job answer the "dream job" question. We ask those who are currently employed whether they received a promotion within the previous six months. Those who received promotions answer the retrospective promotion questions, while those who did not receive promotions answer the prospective promotion questions.

For the prospective happiness questions, the sample sizes are 468 (romantic relationship), 240 (loss of a loved one), 197 (promotion), and 58 (dream job). For the retrospective happiness questions, the sample sizes are 90 (romantic breakups), 60 (loss of a loved one), and 38 (promotion). A feature of the employment

[16] The full text of each survey appears the accompanying "Survey Instruments" document. We conducted this study under human subjects protocol IRB-42264 approved by Stanford University's IRB.

[17] We fielded two waves of the survey on romantic breakups. The text describes Wave 1. Wave 2 was identical, except that we asked a "debiasing" question before the questions about actual and reported feelings: "Imagine that you and the person you're involved with break up within the next week. Using a scale from 1 to 7, where 1 is not at all likely and 7 is very likely, how likely do you think you would be to spend more than an hour a day thinking about the breakup on a typical day two months from now?" As documented in Table 2, this question had no detectable effect on answers to the questions of interest.

survey is that we do not gather the data needed to construct a retrospective measure of self-reported misreporting for the "dream job" event because the potential sample size is too small. Note that the sum of the sample sizes is less than the number of survey participants. The explanation is that some subjects answer neither the retrospective nor the prospective happiness questions—e.g., the romantic breakup survey terminates if the respondent is not in a romantic relationship and has not recently experienced a breakup.

All the prospective happiness questions are worded similarly, as are all the retrospective happiness questions. To investigate the hypothesis concerning order effects articulated in Section 3.1, we randomize the order of the questions concerning feelings and reports of feelings. Using romantic breakups as an example, a subject who is currently in a relationship might first encounter the following prospective question about their anticipated feelings:

> *Imagine that you and the person you're involved with break up within the next week. On a scale of 1 to 7, where 1 is not happy and 7 is very happy, how do you think you would feel on a typical day two months from now?*

The next question would then elicit anticipated reports of feelings:

> *Now that you've told us how you would ACTUALLY feel two months after a breakup, we'd like to know what you think you would SAY IF ASKED on a typical day two months after a breakup to rate your overall happiness.*
>
> *Continue to imagine that you and the person you're involved with break up within the next week. Two months later, you participate in a survey, which asks you how happy you feel on a typical day using a scale from 1 to 7, where 1 is not happy and 7 is very happy. How do you think you would answer that question at that time?*

A subject who is not in a relationship, and who experienced a breakup within the previous six months, might first encounter the following retrospective question about their feelings:

> *On a scale to 1 to 7, where 1 is not happy and 7 is very happy, how happy did you feel on a typical day shortly after your breakup?*

The next question would elicit the feelings they would have reported:

> *Now that you've told us how you ACTUALLY felt after your breakup, we'd like to know what you think you would have SAID IF ASKED on a typical day shortly after your breakup to rate your overall happiness.*
>
> *Suppose you had participated in a survey shortly after your breakup, which asked you how you felt on a typical day using a scale from 1 to 7, where 1 is not happy and 7 is very happy. How would you have answered that question?*

For subjects who encounter the question about reported feelings prior to the question about actual feelings (whether retrospective or prospective), we adjust the wording to reflect the alternative sequencing.

As detailed in Section 3.1, our focus is on the difference between how people say they would (or did) feel, and what they say they would report (or would have reported), with respect to each type of event.

Our survey design allows us to make within-subject comparisons by measuring the difference subject by subject. While our method prescribes posing the expectations question (here, how they would feel) before the reporting question (here, what they say they would report feeling), our design also allows us to assess the effect of reversing the order and to evaluate self-reported misreporting based on cross-subject comparisons of first answers.

As detailed in the next section, the data from our main study reveal a striking difference between self-reported misreporting for the positive and negative events. However, because these events pertain to different domains, it is unclear whether these differences reflect the valence of the events or the domains. The purpose of our supplemental study is to shed light on this question. We gathered the data required to assess self-reported misreporting for two pairs of life events, where the events belonging to each pair pertain to the same domain but have opposite valances. For each currently employed participant, we asked about either receiving a promotion at work (as before) or about being demoted. For each currently unemployed participant, we asked about either being hired for a dream job (as before) or about being forced to accept an unattractive low-paying job. Other aspects of the study's structure were the same as those of the main study.[18] We fielded a single survey on Prolific in June, 2025, with 1004 participants.

4.2. Self-reported misreporting of happiness

We begin by analyzing the data from our main study. Panels A through D of Figure 2 show distributions of responses to questions about happiness conditional on negative events. The figure uses observations for which the expectations question preceded the reporting question, as prescribed. The top two panels, A and B, are for romantic breakups, while the bottom two, C and D, are for death of a loved one. Within each row, the left panel is for prospective events, the right for recalled events. Each panel displays two distributions. The solid bars depict the distribution of how people say they would feel contingent on the event in question, while the hollow bars depict the distribution of what people say they would report. In every case, asking about reported happiness rather than actual happiness shifts the pertinent distribution to the right. In other words, these data exhibit substantial self-reported misreporting. Appendix Figures A.1 and A.2 show that these distributions are qualitatively similar when the order is reversed and when the samples are pooled.

Panels E through G of Figure 2 show distributions of responses to questions about happiness conditional on positive events. The layout is similar to that of Panels A through D, except that the first row pertains to promotions and the second to landing a dream job. For the reasons described in Section 2, there

[18] One exception is that, because we noticed a general decline in the quality of responses on Prolific during the intervening time period, we terminated the supplemental survey for roughly 250 subjects who missed an initial attention check.

is no figure for recalled "dream job" events. Here, we see no evidence of significant systematic self-reported misreporting.

Because we asked each participant about both their contingent happiness and their contingent reported happiness, we can construct measures of self-reported misreporting (SRM) at the subject level. Specifically, for prospective events, SRM equals "expect to feel" minus "expect to report," contingent on the event occurring. For retrospective events, it equals "felt" minus "would have reported" shortly after the event.

Figure 3 displays distributions of SRM. Its structure is identical to that of Figure 1. Each of the seven distributions has a mode of zero. However, the distributions for the negative events in Panels A through D are strikingly left-skewed, indicating that people say they would report feeling happier than they say they would feel. Focusing first on prospective events, we see that for breakups, SRM is strictly positive for only 8% of participants and strictly negative for 55%; for loss of a loved one, these figures are 9% and 48%, respectively. Focusing next on retrospective events, we see that for breakups, SRM is strictly positive for only 25% of participants and strictly negative for 50%; for loss of a loved one, these figures are 20% and 40%, respectively. Moreover, in all cases, strictly positive values of SRM are concentrated at +1 (the smallest possible magnitude). In contrast, the distributions for positive events in Panels E through G exhibit little if any systematic skewness. Appendix Figures A.3 and A.4 show that these distributions are qualitatively similar when the order is reversed and when the samples are pooled.

Table 1 reports regressions that quantify these differences and explore their sensitivity to the order of the questions. The table segregates results for negative and positive events (panels A and B, respectively). We include regressions for both prospective and retrospective events (except the "dream job" event, for which we only have prospective data). For the regressions involving prospective events, we pool the "Expect to Feel" and "Expect to Report" responses. Similarly, for the regressions involving retrospective events, we pool the "Felt" and "Would Have Reported" responses. Accordingly, there are always two observations for each participant.

For each event and time perspective, we report two regressions. All include controls for demographics, a Wave 1 indicator (in the case of romantic breakups), and a constant. The first relates the happiness responses to an indicator for either "Expect to Report" (for prospective events) or "Would Have Reported" (for retrospective events). The corresponding coefficient, which appears in the column labeled "Report," therefore measures the overall average SRM (time negative one). The second regression explores order effects by adding two variables: a "Report Before Feel" indictor (specifying that the question about reported feelings comes first) and an interaction between "Report Before Feel" and "Report." The coefficient of the "Report" dummy measures average SRM (times negative one) when the question about feelings comes first, as our method prescribes. Adding that coefficient to the one for the interaction term gives average SRM (times negative one) when the question about reported feelings comes first (which is not preferred). The regression also allows one to infer SRM only from first responses using cross-subject variation (which is likewise not preferred): to obtain the estimate, we simply add the coefficients of "Report," "Report Before Feel," and the interaction, and then change signs. All standard errors are clustered at the individual level.

According to equation (1), for prospective romantic breakups, the overall mean for SRM is -0.58 ($SE = 0.06$), indicating that people think they will report feeling substantially happier than they say they will feel. To put this figure in perspective, it represents 39.7% of a standard deviation for the distribution of prospective reported happiness for this event. We reject the absence of a difference with high confidence ($p < 0.001$). The coefficients in equation (2) reflect significant order effects: as conjectured, the magnitude of average SRM is larger (-0.77, $SE = 0.09$) when the questions appear in the prescribed order (the expectations question first), and smaller (-0.38, $SE = 0.08$) when they appear in the opposite order (the reporting question first). In both cases we continue to reject the absence of a difference with high confidence. The cross-subject first-response estimate of SRM is -0.73 ($SE = 0.14$). A close inspection of the coefficients reveals that asking the reporting question first increases reported happiness for the expectations question (by 0.34) without materially changing the response to the reporting question. In other words, the higher answer to the reporting question exerts an upward pull on the expectations answer (as conjectured) but does not eliminate the difference.

Equation (3) shows that the overall mean value of SRM for retrospective breakups, -0.61 ($SE = 0.17$), is similar to that for prospective breakups. Once again, we reject the absence of an effect with high confidence ($p < 0.01$). The point estimates for equation (4) reflect the same pattern of order effects: the magnitude of average SRM is larger (-0.77, $SE = 0.29$) when the questions appear in the prescribed order (the recollection question first), and smaller (-0.35, $SE = 0.20$) when they appear in the opposite order (the reporting question first). The smaller sample size ($N = 180$ versus 936) reduces precision, rendering the interaction effect statistically insignificant, though economically large. The cross-subject first-response estimate of SRM, -0.29 ($SE = 0.35$), is based on an even smaller sample (56 observations for the recollection question and 44 observations for the reporting question) and lacks statistical significance.

For prospective loss of a loved one, equation (5) shows that the overall mean of SRM is similar, -0.49 ($SE = 0.07$), representing 37.8% of a standard deviation for the distribution of reported happiness for this event. Interestingly, in equation (6), we see no evidence that the order of presentation affects our estimate of SRM (although the standard error of the interaction term is large). The cross-subject first-response estimate of SRM is -0.81 ($SE = 0.17$). Patterns for retrospective questions are similar (equations (7) and (8)), but the sample size is even smaller than for romantic breakups ($N = 120$), so the coefficients are not statistically significant. However, the cross-subject first-response estimate of SRM is -0.80 ($\sigma = 0.30$).

Consistent with our discussion of Panels E to G of Figures 1 and 2, we mostly obtain null effects for positive events. The overall means for SRM cluster around zero, and the coefficients generally lack statistical significance. The one exception is that we find significant order effects for the prospective questions about job promotions (equation (10)). Indeed, when the questions about prospective promotions appear in the prescribed order, we find that SRM is 0.31 ($SE = 0.09$), indicating that people expect to report being *less* happy than they feel. A change in the order flips the sign of the difference; overall, the mean difference is 0.02 ($SE = 0.07$).

Because the positive and negative events considered in the main study fall within different life-experience domains, we cannot say whether the stark differences between the results in Panels A and B of Table 1 are attributable to valance or domain. We address this issue by analyzing data from the supplemental study. As explained in Section 4, these data cover two pairs of life events, where the events belonging to each pair pertain to the same domain but have opposite valances. Analyses of these data appear in Panels C and D of Table 1.

Focusing on those who are currently employed, we see from equations (15) and (19) that the overall average SRM is -0.49 ($SE = 0.06$) conditional on being demoted, compared with -0.22 ($SE = 0.05$) conditional on receiving a promotion. The difference (0.27) is highly statistically significant ($SE = 0.08$, $p < 0.01$). Similarly, focusing on those who are currently unemployed, we see from equations (17) and (21) that the overall average SRM is -0.83 ($SE = 0.15$) conditional on being forced to accept an undesirable job, compared with 0.02 ($SE = 0.15$) conditional on landing a dream job. Once again, the difference (0.85) is highly statistically significant ($SE = 0.08, p < 0.001$). We find qualitatively similar patterns when the questions appear in the prescribed order. From equations (16) and (20), we see that the SRM is -0.53 ($SE = 0.08$) conditional on being demoted versus -0.07 ($SE = 0.07$) conditional on being promoted. Similarly, from equations (18) and (21), we see that the SRM is -0.87 ($SE = 0.18$) conditional on being forced to accept an undesirable job versus 0.39 ($SE = 0.20$) conditional on landing a dream job. Thus, evidence from our supplemental study confirms that valance plays a significant role in determining the degree of self-reported misreporting.[19]

Taken as whole, the findings reported in this section have troubling implications for the use of happiness data. Suppose, for example, that a negative event occurring simultaneously with the receipt of

[19] Comparing equations (9), (10), (13), and (14) to, respectively, equations (19) through (22), we see that results based on our supplemental study are qualitatively similar to the results from our initial study for promotions and dream jobs. Most of the quantitative differences lack statistical significance. The statistically significant differences may reflect changes in the population of Prolific participants or in economic conditions (and hence attitudes toward employment) over the intervening two years.

$100 leaves reported happiness unchanged. One might be tempted to conclude that the compensating variation for the event is $100. But if, as our results suggest, people understate the impact on happiness of certain events while correctly stating the impact of others (including the receipt of money), the true compensating variation may be greater than $100. More generally, our results suggest that the magnitude of the reporting bias may be related to the specific combination of salient events the respondent has recently experienced. In other words, this bias does not necessarily preserve the ordering of outcomes, as would a uniform bias or simple compression.

4.3. Connecting SRM to reports of contemporaneous happiness

Do people who say they would misreport expected happiness conditional on a specified event contemporaneously misreport their experienced happiness, as our method assumes? As discussed in Section 3.1, we can draw inferences about actual misreporting from statements about expected misreporting by assuming either that people correctly anticipate what they would report under specified contingencies, or that one form of domain-specific dishonesty potentially implies another. Here, we provide more direct evidence.

Even though we cannot document contemporaneous misreporting directly (because true contemporaneous happiness is unobserved), we can still explore whether the mechanisms that drive misreporting of expected conditional happiness also infect reports of current happiness. Specifically, we investigate whether people who expect to exaggerate their happiness when asked about a future contingency also report greater happiness when asked about their current state of mind.

Our empirical strategy is to estimate regressions of the following form:

$$q_i = \alpha + \beta q_{ix}^h + \gamma SRM_{ix} + X_i\delta + \varepsilon_{ix},$$

where q_i is the current happiness subject i reports, q_{ix}^h is the happiness i says they expect to feel after event x, SRM_{ix} is i's self-reported misreporting for event x, X_i is a vector of personal characteristics, and ε_i is a disturbance term. In other words, for any given event x, the regression describes the relationship between happiness and self-reported misreporting conditional on the happiness i reportedly expects to feel after event x. We expect to find a relationship between q_i and q_{ix}^h because q_{ix}^h proxies for i's general level of happiness as well as for the cardinalization i adopts when reporting happiness. However, conditional on q_{ix}^h, it is reasonable to assume that SRM_{ix}^e is unrelated to i's baseline happiness or cardinalization: it is simply a statement concerning the degree to which i expects to depart from q_{ix}^h when reporting happiness under the same hypothetical. Therefore, we would only expect to find $\gamma < 0$ if the mechanisms that drive self-reported misreporting also affect the accuracy with which respondents report current happiness—for example, if the people who say they would exaggerate q_{ix}^h after specifying it also exaggerate current happiness. Because this reasoning assumes that the expectations question establishes q_{ix}^h prior to the reporting question, we estimate the regressions using the subset of observations for which these questions appear in the prescribed order.[20]

Results appear in Table 2. There are five regressions, one for each of the four prospective events covered in the main study, and one pooled regression that also includes dummies for the event types. As expected, β is strictly positive and statistically significant in all five regressions. More importantly, γ is strictly negative in all five regressions. Although it lacks statistical significance in three of the event-specific regressions, it is highly statistically significant for promotions and, critically, in the pooled regression. These findings are consistent with the hypothesis that self-reported misreporting reflects tendencies that infect contemporaneous reports of happiness.

4.4. Implications for projection bias

The final step in our analysis is to illustrate how SRM calculations can impact the conclusions one draws from happiness data. Specifically, we use our results to evaluate the extent to which estimates of projection bias that employ happiness data (as in Gilbert et al., 1998, and Loewenstein et al., 2003) reflect misreporting of experienced happiness once events occur.

The term "projection bias" refers to a tendency for people to exaggerate the resemblance between their current and future tastes. A widely used strategy for measuring projection bias involves comparing happiness experienced after a positive or negative event with prospective expectations of happiness contingent on that event occurring. The standard finding is that people believe negative events will lead to lower happiness, and positive events to greater happiness, than is truly the case. The SRM method invites a reevaluation of the evidence on projection bias partly because the survey questions we pose concerning expected happiness contingent on future events resemble those used in that application. Our findings raise the possibility that evidence of projection bias may be attributable, in whole or in part, to systematic misreporting of happiness. Specifically, if people exaggerate their happiness when responding to survey questions after bad events, they will artificially appear to be more adaptable than they expect.

To gauge the potential significance of misreporting, we will assume provisionally that self-reported misreporting provides an accurate (or at least conservative) measure of the degree to which people misreport happiness. Under that assumption, we can estimate *ex post* happiness following some specified event by adding SRM to reported happiness. We then obtain an alternative estimate of projection bias by subtracting predicted happiness from this adjusted measure of *ex post* happiness.

[20] Results are stronger when we include observations for which the reporting question preceded the expectations questions. However, we do not report those results because, as noted previously, asking the reporting question first potentially cross-contaminates the two answers, thereby obscuring the interpretations of β and γ.

Table 3 uses our data to construct alternative estimates of projection bias for three of the four events we consider in the main survey (romantic breakups, losses of loved ones, and promotions).[21] It displays the means of (1) ex post happiness, reported by those who have recently experienced the event,[22] (2) anticipated happiness after experiencing the event, reported by those who have not recently experienced it, and (3) the standard measure of projection bias (the difference between (1) and (2)). We find strong evidence of projection bias for the negative events we consider. The difference between the means is 1.46 ($SE = 0.17$) for romantic breakups and 2.23 ($SE = 0.17$) for loss of a loved one. The usual interpretation of such findings is that those who experience these events are happier than they anticipated. We do not find evidence of projection bias for the positive event (promotion). Contrary to the standard hypothesis that people feel less happy than they anticipated after positive events, we find that the difference in means is *positive* (0.13). However, it is small in magnitude and statistically insignificant ($SE = 0.21$).

To adjust these measures of projection bias for self-reported misreporting, we use the estimates of SRM derived from the column labeled "Report" in the even-numbered equations of Table 1, which are based on observations for which the questions appeared in the prescribed order. For prospective romantic breakups, our preferred estimate of average SRM is -0.77, so the adjustment reduces measured projection bias by 53%. For loss of a loved one, our preferred estimate of average SRM is -0.52, so the adjustment reduces measured projection bias by 24%. For promotions, our preferred estimate of SRM is 0.31 based on data from our main study and 0.07 based on data from our supplemental study. Because these numbers are positive, the adjusted estimates are even less consistent with the standard projection bias hypothesis than the unadjusted estimate. However, the adjustment based on the supplemental sample is small and statistically insignificant.

For the negative events we consider, these results imply that a significant fraction of estimated projection bias is attributable to misreporting. However, our adjustments fall well short of erasing evidence of bias entirely. It is, therefore, essential to keep four considerations in mind. First, for the reasons discussed in Section 3.1, SRM may well be a lower bound on the bias in contemporaneous reports of happiness. In that case, our modified measure of projection bias serves as an upper bound rather than as a point estimate. Second, as emphasized in Section 3.1, questions about happiness contingent on future events invariably

[21] As explained in Section 4.1, we do not gather the data needed to construct a retrospective measure of self-reported misreporting for the "dream job" event because the potential sample size is too small. The same issue renders this event unsuitable for studying projection bias.

[22] As explained in Section 4.1, our surveys identify subjects who have experienced the events in question within the previous six months. We also ask how many months prior to the survey the event occurred. Consequently, we could in principle identify subjects who experienced the event shortly before the survey, just as the prospective question asks. However, the resulting samples would be extremely small. Analyses of the data reveal no relationships between happiness responses and the time elapsed since the occurrence of the event. We therefore report average happiness for all respondents who experienced the event within the previous six months.

describe incomplete hypothetical scenarios. Estimates of projection bias assume that, when the survey question specifies the event x^h, people report the expected value of q considering all compatible realizations of z and using the Likert scale cardinalization, given their beliefs about the conditional distribution of z and the mapping from the completely specified event (x, z) to q. In practice, they may answer by completing the hypothetical with a value of z that allows them to give an answer they prefer without lying, or the one they consider most likely. Alternatively, they may assess expected values based on some alternative cardinalization and then convert to the Likert scale. Or they may simply be wrong about the distribution of z. Such considerations prevent one from attributing differences between realized happiness and statements about anticipated happiness to misunderstandings of the mapping from completely specified events to happiness (i.e., Q), which is the essence of projection bias. Third, because happiness is measured on a unitless scale, respondents may renormalize the scale when their circumstances change. If they always use the same numerical value to indicate happiness on a (locally) typical day, they will appear more adaptable than they actually are. Fourth, when the respondent can influence the likelihood with which an event occurs, selection effects can confound estimates of projection bias. For example, people who are in less satisfying romantic relationships may be more likely to experience breakups and happier when breakups occur. Even if expectations are accurate, anticipated happiness conditional on breakups for people in relationships would then be lower than average happiness after breakups. Accordingly, even if misreporting does not fully explain the evidence commonly cited in favor of projection bias, it could potentially do so in combination with other factors.

To be clear, the literature on projection bias also marshals other types of evidence. For example, the current weather appears to exert disproportionate influence over the purchase and retention of durable goods (see, e.g., Conlin, O'Donoghue, and Vogelsang, 2007, and Busse, Pope, Pope, and Silva-Risso, 2015). However, these findings have other possible explanations—for example, that people do not have well-defined preferences, and that cues impacting the salience of an aspect of experience may influence the weight it receives during process of "preference construction."

5. Conclusions

In this paper, we have introduced a method for validating survey responses in contexts where there are no ground-truth benchmarks. First, we direct people to envision a hypothetical scenario in which the quantity of interest would be q. One can either fix q as part of the scenario's description or, to ensure that respondents find the hypothetical plausible, ask them to indicate the level of q they think would prevail under the specified conditions. Then we ask them to indicate the response they think they would give if asked to report q on a survey under precisely the same conditions. Self-reported misreporting is the difference between q and the value they say they would anticipate reporting.

We have clarified the inferences one can draw about actual misreporting from SRM under various combinations of assumptions. As a general matter, non-zero SRM implies that the respondent either expects to misreport q or misrepresents the value they expect to report (or both). Either possibility calls the respondent's reliability into question, thereby undermining confidence in reported values of q. Stronger conclusions follow under various combinations of additional assumptions. For example, we have identified assumptions under which one can treat SMR as a lower bound on the respondent's actual expectation concerning misreporting, as an estimate of that expectation, as a lower bound on actual misreporting, and as an estimate of actual misreporting. We also have also explained why it may be reasonable to think these various assumptions are satisfied.

To validate our validation method, we have conducted an experiment based on an established paradigm for measuring the extent to which people lie when only they know the truth. In this setting, we have found that, on average, SRM coincides almost exactly with actual misreporting. This finding lends credibility to the assumptions that justify SRM as a measure of misreporting, rather than merely as a bound or diagnostic.

Applying the method to happiness data, we have concluded that SRM is substantial and varies systematically across experiences. We have also provided evidence that the mechanisms driving expected misreporting of conditional happiness infect reports of current happiness. Specifically, people who say they expect to exaggerate their happiness when asked about a future continency also report greater happiness when asked about their current state of mind. Finally, we illustrate the significance of our findings by showing that evidence of projection bias is attributable, in significant part, to misreporting. Collectively, these findings suggest that considerable caution is warranted when interpreting studies that attempt to infer how conditions affect welfare by analyzing measures of self-reported well-being.

We emphasized at the outset that our validation method is potentially applicable to other types of survey data for which ground-truth benchmarks are unavailable. We listed several examples in Section 1, some of which we hope to explore in subsequent research. For any application, the key is to formulate suitable hypothetical scenarios.

As an illustration, consider the problem of assigning monetary values to environmental goods. The contingent valuation method proceeds by eliciting stated willingness-to-pay for a hypothetical course of action that targets public resources, such as cleaning up oil spills. Efforts to validate these measures have focused on market goods, but such goods may activate somewhat different motives than public goods. Our method provides a viable alternative. To apply it, one would field a survey that describes a prospective hypothetical event (e.g., a future oil spill), asks participants what they would be willing to pay in that contingency for public action designed to mitigate the adverse effects of the event, and then asks them what they would *say* they would be willing to pay if asked on a survey in the same contingency. Systematic differences between statements about what they would be willing to pay if the event occurred, and what they would say they would be willing to pay, would call the reliability of hypothetical statements about willingness-to-pay to remediate actual events into question.

Some applications may be even simpler. To evaluate the accuracy of responses to questions about the frequency of exercise, one could ask people to state the frequency they think they would report if asked on a survey, contingent on some hypothetical "true" frequency. To evaluate the accuracy of responses to questions about a sensitive financial quantity such as income or savings, one could ask people to state the level they think they would report if asked on a survey, contingent on a hypothetical "true" level. To evaluate the accuracy of answers to questions about the consumption of a "sin" good such as alcohol, one could specify a hypothetical contingency involving a given level of average weekly consumption, and then ask people to say what they think they would report in that hypothetical scenario if asked about their consumption level on a survey. In each application, one would need to evaluate whether it is reasonable to specify the hypothetical scenarios (the basic method) or elicit them (the extended method). In principle, there could also be a variety of context-specific issues. Still, the general method may prove broadly applicable.

References

Benoit, J.-P., & Dubra, J. (2011), “Apparent Overconfidence,” *Econometrica* 79(5), 1591-1625.

Benjamin, D.J., O. Heffetz, M.S. Kimball, and A. Rees-Jones (2012), “What Do You Think Would Make You Happier? What Do You Think You Would Choose?” *American Economic Review* 102(5), 2083-2110).

Benjamin, D.J., O. Heffetz, M.S. Kimball, and N. Szembrot (2014), “Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preferences,” *American Economic Review* 10(9), 2698-2735.

Bentham, Jeremy (2023). *An Introduction to the Principles of Morals and Legislation*. Oxford University Press, Oxford.

Bernheim, B.D., Taubinsky, D., and Kim (forthcoming), “Welfare and the Act of Choosing,” *Journal of Political Economy*.

Bond, T. N., & Lang, K. (2019), “The sad truth about happiness scales,” *Journal of Political Economy* 127(4), 1629-1640.

Busse, M. R., Pope, D. G., Pope, J. C., and Silva-Risso, J. (2007), “The Psychological Effect of Weather on Car Purchases,” *Quarterly Journal of Economics* 130(1), 371-414.

Conlin, M., O’Donoghue, T., and Vogelsang, T. J. (2015), “Projection Bias in Catalog Orders,” American Economic Review 97(4), 1217-1249.

Davidson, R. J. (2004), “Well-being and affective style: neural substrates and biobehavioral correlates,” *Philosophical Transactions, Biological Sciences* 359, 1395-1341.

Diener, E., & Oishi, S. (2004), “Are Scandinavians Happier than Asians? Issues in comparing Nations on Subjective Well-Being,” in F. Columbus (ed.), *Asian Economic and Political Issues*, 1-25.

Fischbacher, U., and Föllmi-Heusi, F. (2013), “Lies in Disguise—An Experimental Study on Cheating,” *Journal of the European Economic Association* 11 (3), 525-547.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., and Wheatley, T. P., “Immune Neglect: A Source of Durability Bias in Affective Forecasting,” *Journal of Personality and Social Psychology* 75, 617-638.

Goldings, H. J. (1954), “On the avowal and projection of happiness,” *Journal of Personality* 23, 30-47.

Hausman, Daniel M. (2012). *Preference, Value, Choice, and Welfare*. Cambridge University Press, Cambridge.

Helliwell, J. F., Layard, R., & Sachs, J. (2014). *World Happiness Report 2013*. The Earth Institute, Columbia University, New York.

Kagan, Shelly (1998). *Normative Ethics*. Westview Press, Boulder, Colorado.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., Stone, A. (2004), “Toward national well-being accounts,” *American Economic Review* 94 (2), 429–434.

Larsen, R. J., & Fredrickson, B. L. (1999), “Measurement Issues in Emotion Research,” in Kahneman, D., Diener, E., and Schwarz, N., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation.

Liu, S., & Netzer, N. (2023), “Happy times: Identification from ordered response data,” University of Zurich, Department of Economics, Working Paper.

Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003), “Projection bias in predicting future utility,” *Quarterly Journal of Economics*, 1209-1248.

Mill, John Stuart (2012). *Utilitarianism*. Renaissance Classics.

Parfit, Derek, 1984. *Reasons and Persons*. Oxford University Press, Oxford.

Santos Pinto, L., & Sobel, J. (2005), “A model of positive self-image in subjective assessments,” *American Economic Review* 95(5), 1386-1402.

Stutzer, A., & B. Frey (2008), “Stress that doesn’t pay: the commuting paradox,” *Scandinavian Journal of Economics* 110(2), 339-366.

Van Praag, B. M. S., & Ferrer-i-Carbonell, A. (2004). *Happiness Quantified: A Satisfaction Calculus Approach*. Oxford, U.K.: Oxford University Press.

Veenhoven, R. (1984). *Conditions of happiness*. Dordrecht, Netherlands: Reidel.

Veenhoven, R. (2008), “Comparability of happiness across nations,” *School of Sociology and Social Work Journal* 104, 211-234.

Venhoven, R. (2015), “Social conditions for human happiness: A review of research,” *International Journal of Psychology* 50(5), 379-391.

Figure 1: CDFs for Number of Heads (Actual, Reported, and Predicted Reports)

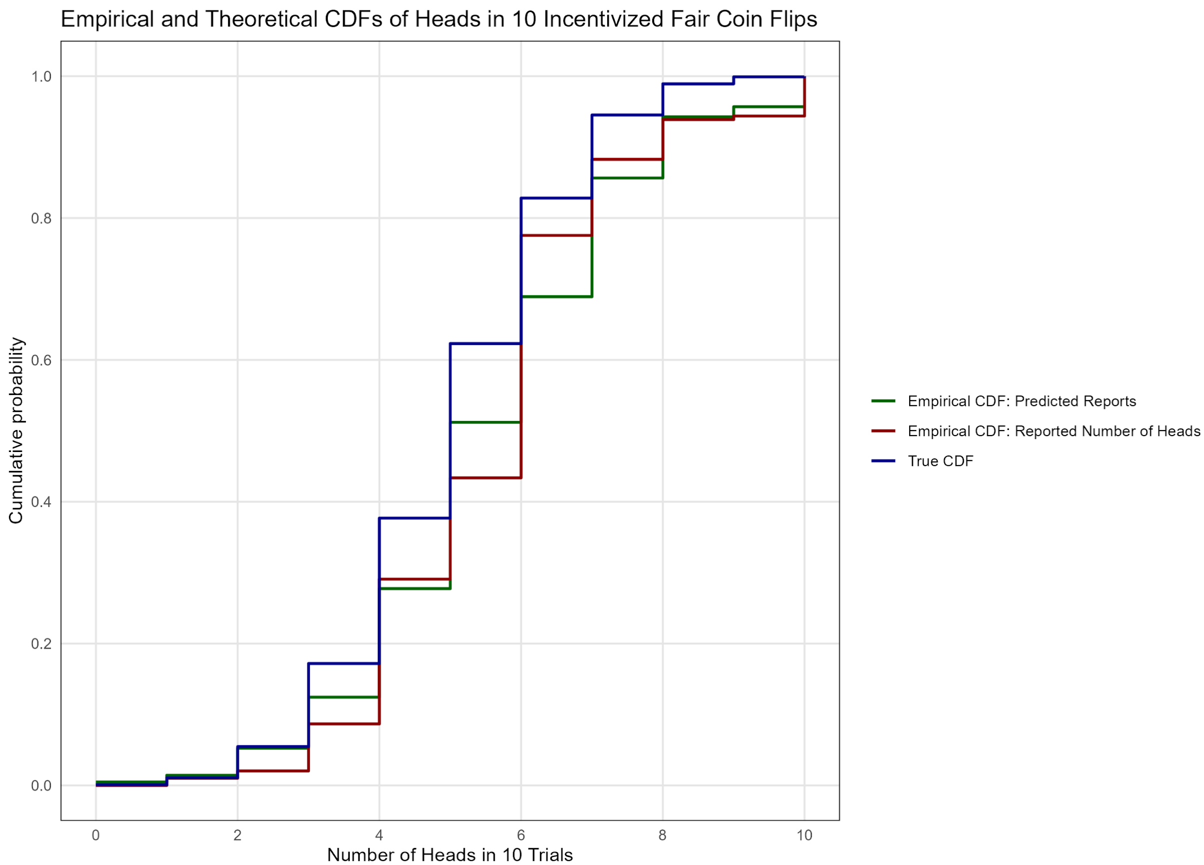
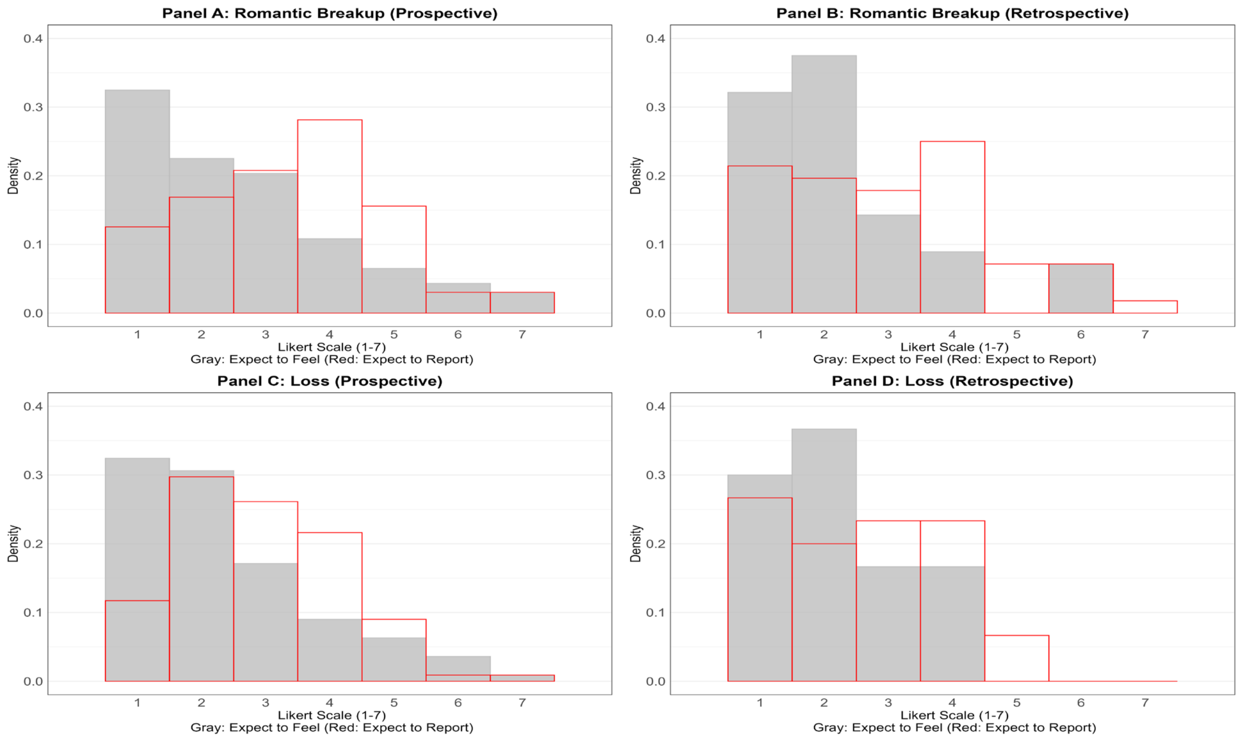


Figure 2: Expected Happiness versus Expected Reports of Happiness, “Feel” Before “Report”

Negative Events: Panels A through D



Positives Events: Panels E through G

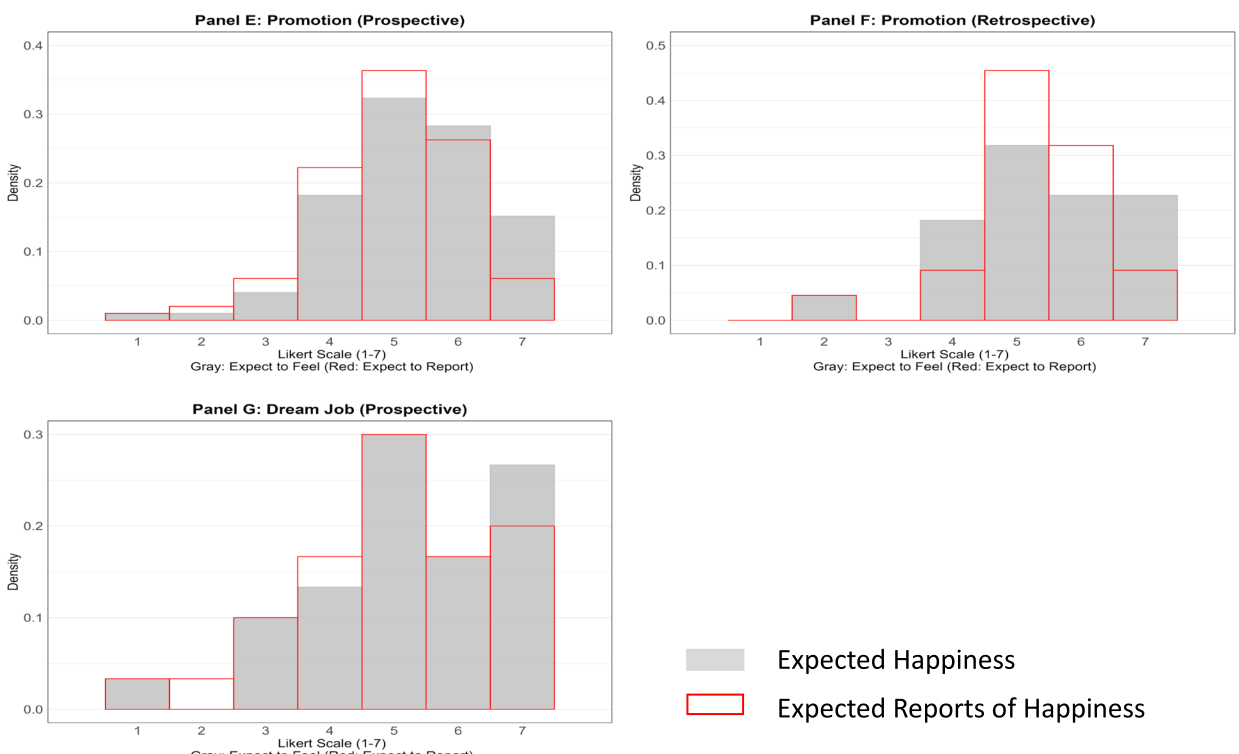
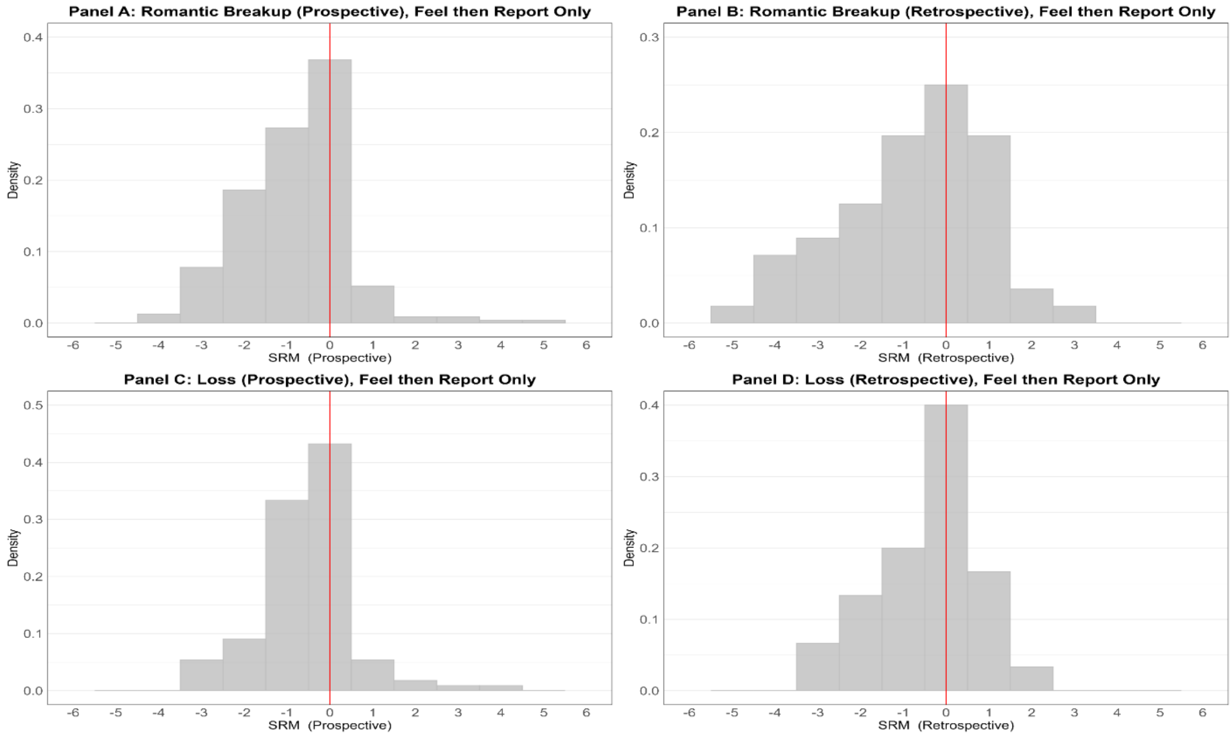


Figure 3: Distributions for Self-Reported Misreporting (within subject), “Feel” Before “Report”

Negative Events: Panels A through D



Positive Events: Panels E through G

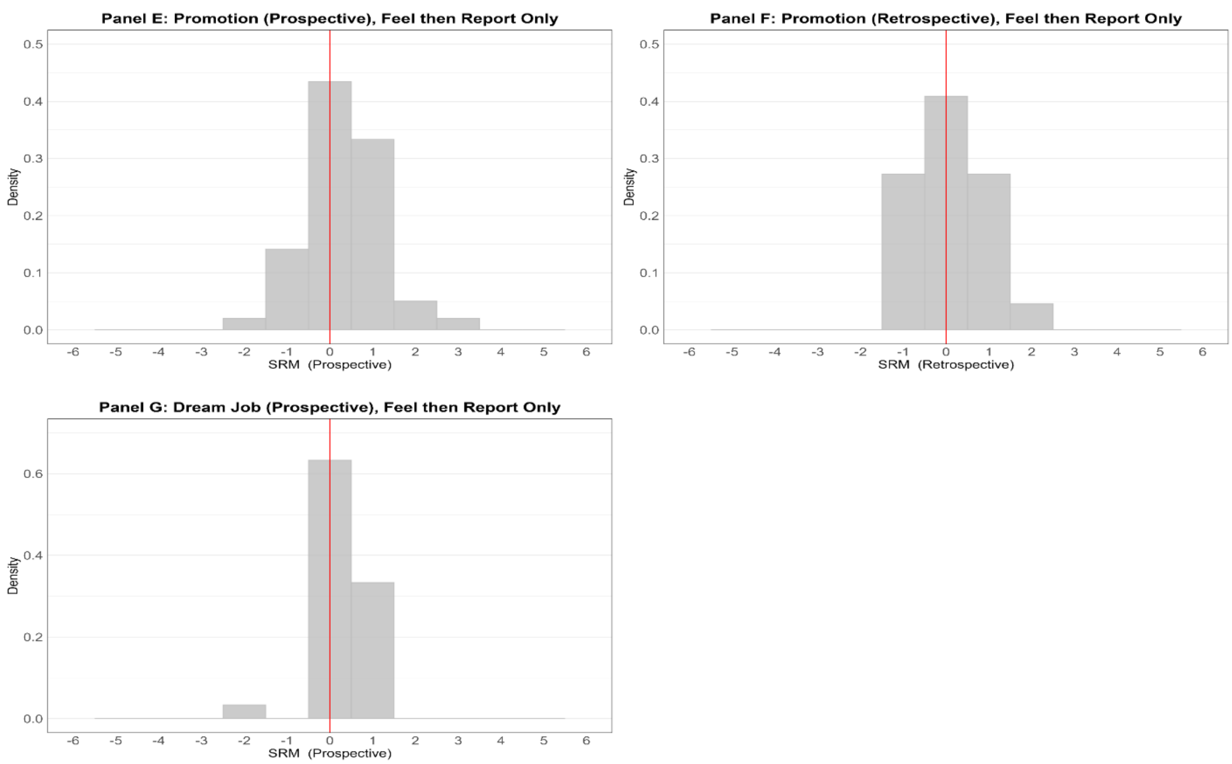


Table 1: Pooled Regressions for Expected Happiness and Expected Reports of Happiness

	Event	Temporal Perspective	Report	ReportBeforeFeel × Report	ReportBeforeFeel	Constant	Observations
Panel A: Negative Events (Main Surveys)							
(1)	Breakup	Prospective	0.575*** (0.059)			2.938*** (0.133)	936
(2)	Breakup	Prospective	0.771*** (0.089)	-0.387*** (0.118)	0.335** (0.143)	2.766*** (0.159)	936
(3)	Breakup	Retrospective	0.611*** (0.166)			2.446*** (0.250)	180
(4)	Breakup	Retrospective	0.768*** (0.236)	-0.415 (0.311)	-0.068 (0.318)	2.468*** (0.242)	180
(5)	Loss of Loved One	Prospective	0.488*** (0.068)			2.563*** (0.163)	480
(6)	Loss of Loved One	Prospective	0.523*** (0.107)	-0.065 (0.137)	0.347** (0.171)	2.35*** (0.196)	480
(7)	Loss of Loved One	Retrospective	0.367*** (0.136)			2.668*** (0.297)	120
(8)	Loss of Loved One	Retrospective	0.433 (0.226)	-0.133 (0.274)	0.495 (0.329)	2.449*** (0.313)	120
Panel B: Positive Events (Main Surveys)							
(9)	Promotion	Prospective	0.015 (0.065)			4.739*** (0.139)	394
(10)	Promotion	Prospective	-0.313*** (0.093)	0.660*** (0.122)	-0.600*** (0.164)	5.036*** (0.162)	394
(11)	Promotion	Retrospective	0.000 (0.149)			5.431*** (0.329)	76
(12)	Promotion	Retrospective	-0.091 (0.19)	0.216 (0.308)	0.211 (0.406)	5.331*** (0.355)	76
(13)	Dream Job	Prospective	0.069 (0.117)			4.938*** (0.348)	116
(14)	Dream Job	Prospective	-0.267** (0.118)	0.695*** (0.219)	0.02 (0.395)	4.893*** (0.469)	116
Panel C: Negative Events (Supplemental Survey)							
(15)	Demotion	Prospective	0.491*** (0.063)			2.247*** (0.077)	758
(16)	Demotion	Prospective	0.525*** (0.082)	-0.073 (0.127)	0.021 (0.122)	2.237*** (0.097)	758
(17)	Undesirable Job	Prospective	0.831*** (0.146)			1.815*** (0.134)	130
(18)	Undesirable Job	Prospective	0.865*** (0.176)	-0.079 (0.305)	-0.098 (0.214)	1.865*** (0.174)	130
Panel D: Positive Events (Supplemental Survey)							
(19)	Promotion	Prospective	0.224*** (0.052)			5.448*** (0.075)	820
(20)	Promotion	Prospective	-0.065 (0.070)	0.611*** (0.099)	-0.347*** (0.113)	5.611*** (0.088)	820
(21)	Dream Job	Prospective	-0.016 (0.153)			5.420*** (0.199)	128
(22)	Dream Job	Prospective	-0.387 (0.202)	0.720** (0.294)	-0.479 (0.279)	5.665*** (0.202)	128

Note: all regressions cluster standard errors at the individual level and control for gender; Panels A & B additionally control for survey wave.

Table 2: Regressions of Current Happiness on Contingent Expectations

Table 2: Current Surveyed Happiness: (Feel-Then-Report Sample)					
	(1)	(2)	(3)	(4)	(5)
Event	**Breakup**	**Death of a loved one**	**Promotion**	**Dream Job**	**All Events**
Time Perspective	Prospective	Prospective	Prospective	Prospective	Prospective
Feel (β)	0.114**	0.429***	0.587***	0.672***	0.289***
	(0.045)	(0.088)	(0.081)	(0.118)	(0.035)
SRM (γ)	-0.020	-0.167	-0.303***	-0.232	-0.119***
	(0.053)	(0.110)	(0.102)	(0.279)	(0.044)
Gender	-0.135	0.145	0.368**	-0.498	0.056
	(0.125)	(0.208)	(0.169)	(0.356)	(0.091)
Religiosity	0.180***	0.159***	0.026	0.159*	0.156***
	(0.033)	(0.050)	(0.051)	(0.089)	(0.024)
Wave2 x Breakup	0.050				0.031
	(0.123)				(0.130)
Breakup					3.425***
					(0.165)
Death					3.373***
					(0.157)
Promotion					2.783***
					(0.214)
Dream Job					1.815***
					(0.262)
Observations	231	111	99	30	471

Note: all regressions estimated with an intercept except for regression 5.

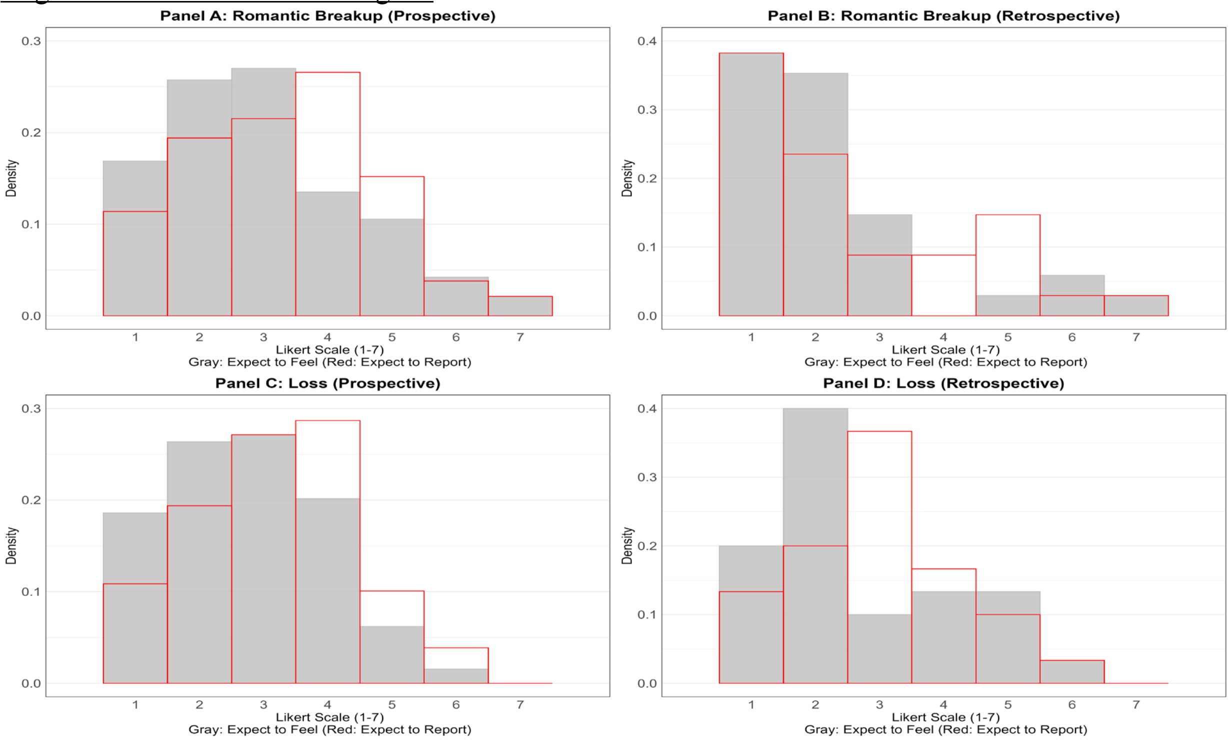
Table 3: Measured Projection Bias

		Romantic Breakup	Death of Loved One	Promotion
Ex Post Report	Mean	4.250	4.817	5.106
	SD	1.195	1.097	1.147
	N	90	60	38
Report of Ex Ante Anticipation	Mean	2.791	2.583	4.980
	SD	1.563	1.345	1.220
	N	468	240	197
Ex Post Report - Report of Ex Ante Anticipation	Mean	1.459	2.233	0.127
	SE	0.145	0.166	0.205
	N1	90	60	38
	N2	468	240	197

Appendix: Additional Results

Figure A.1: Expected Happiness versus Expected Reports of Happiness, “Report” Before “Feel”

Negative events: Panels A through D



Positive events: Panels E through G

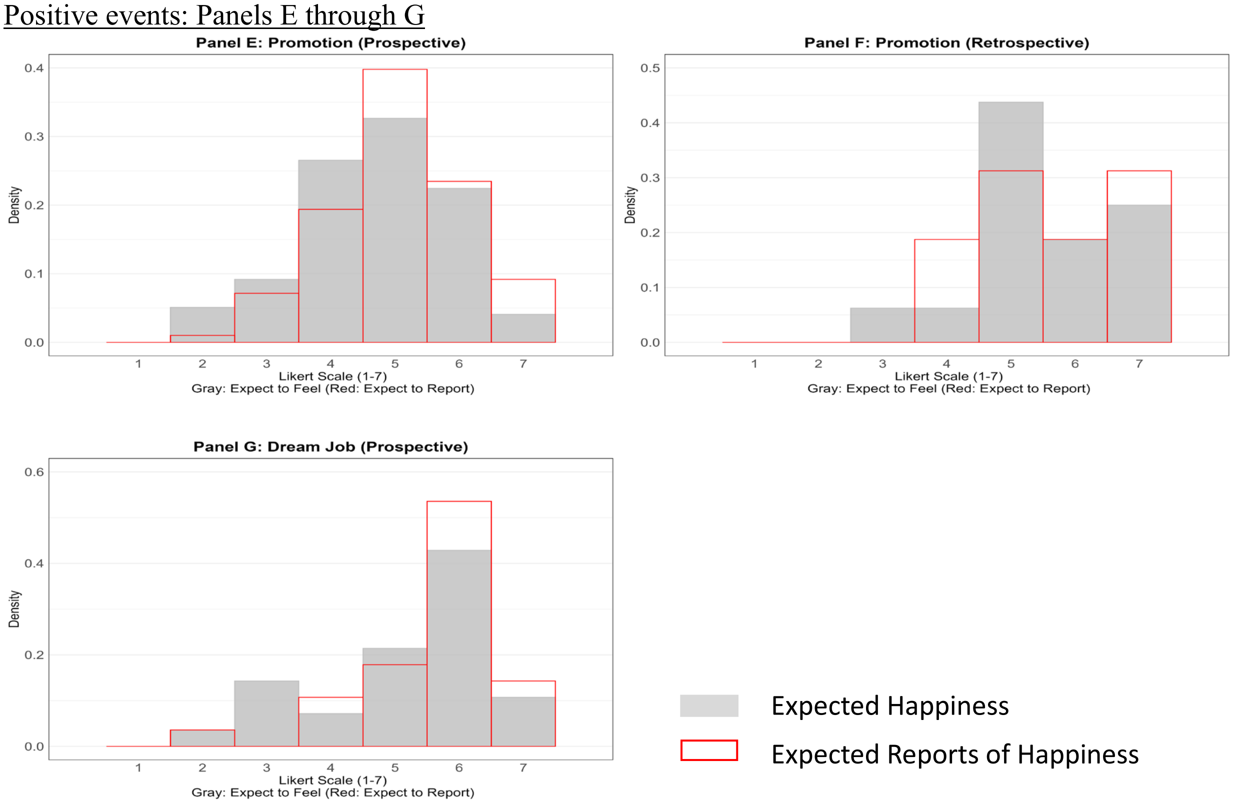
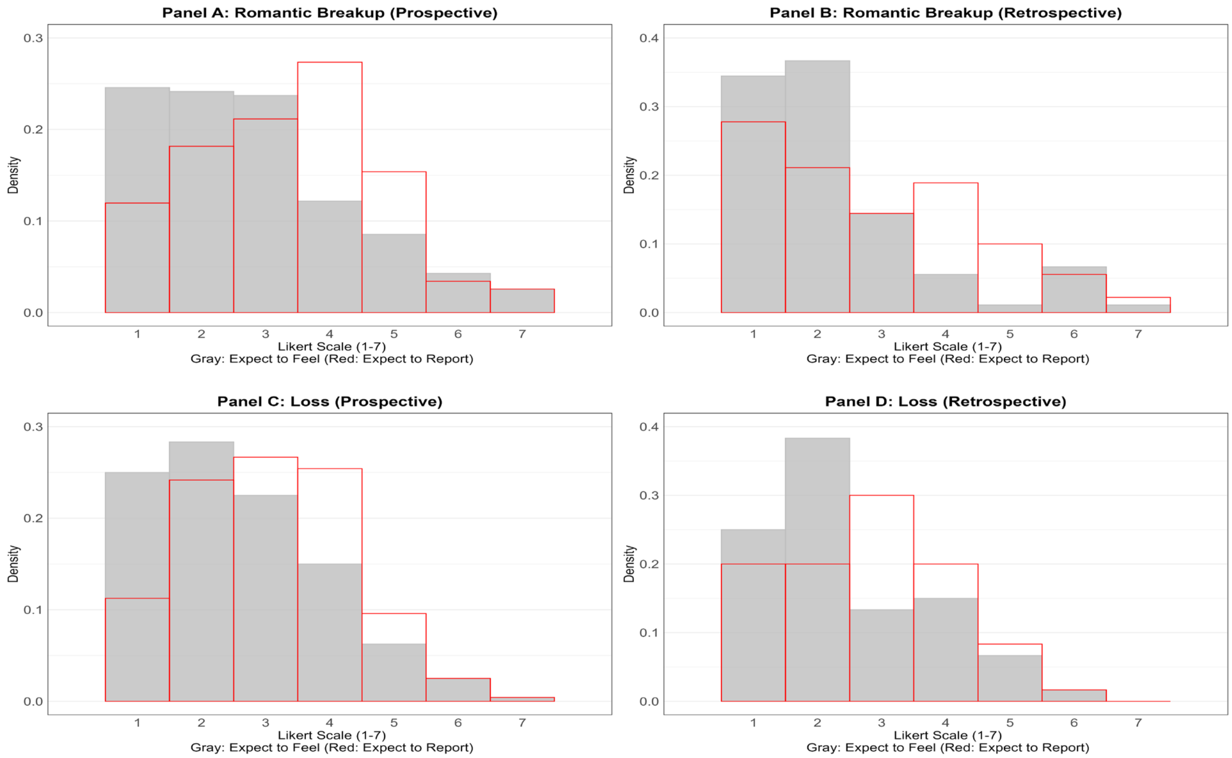


Figure A.2: Expected Happiness versus Expected Reports of Happiness, Pooled

Negative events: Panels A through D



Positive events: Panels E through G

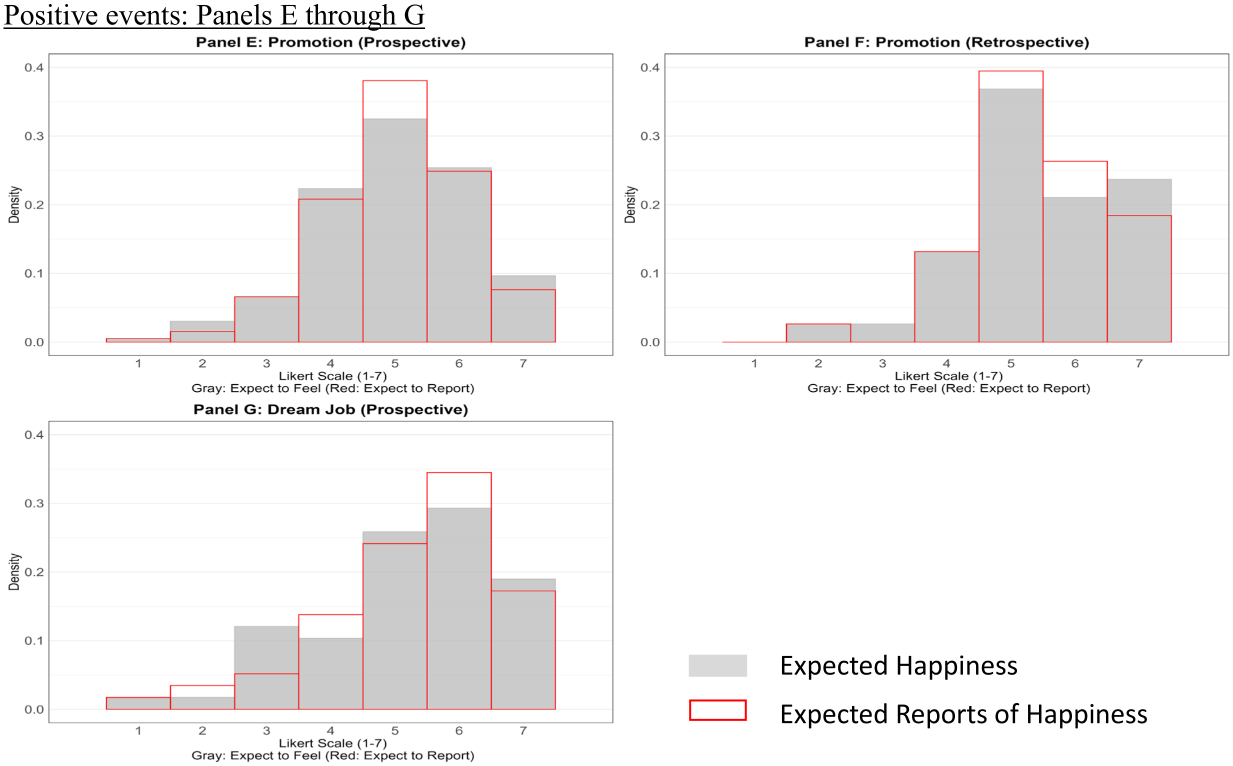
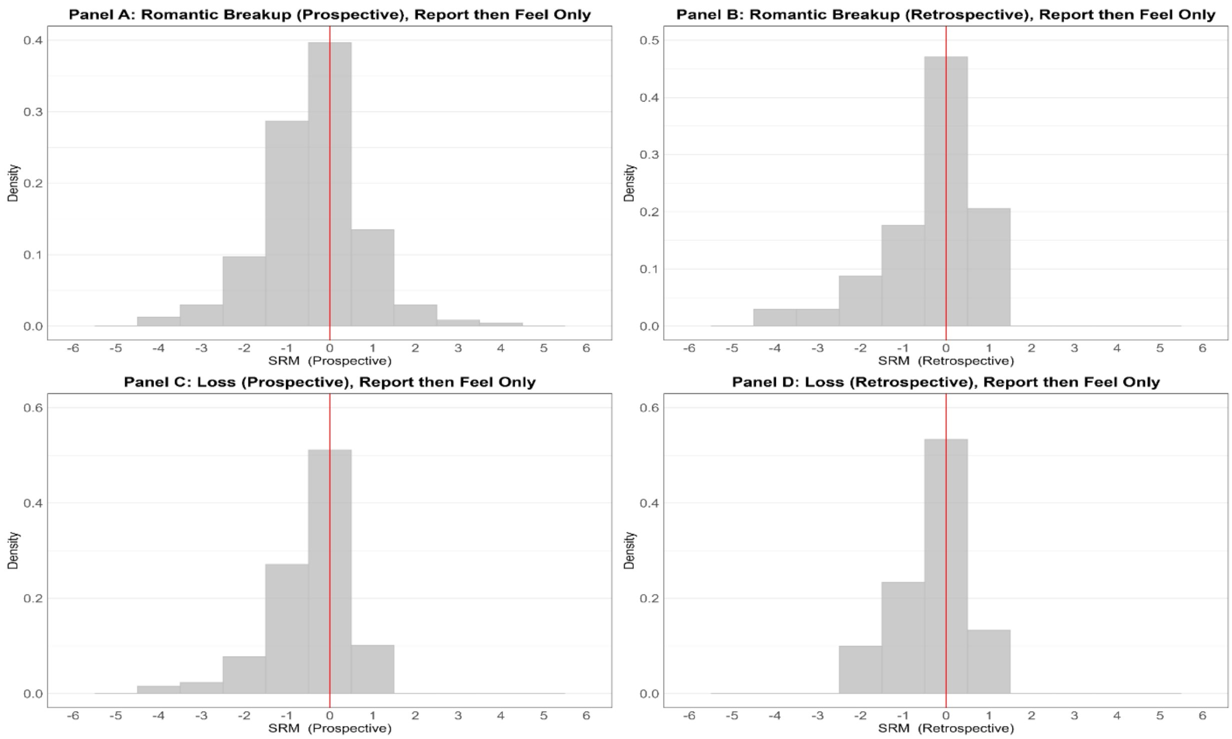


Figure A.3: Distributions for Self-Reported Misreporting (within subject), “Report” Before “Feel”

Negative Events: Panels A through D



Positive Events: Panels E through G

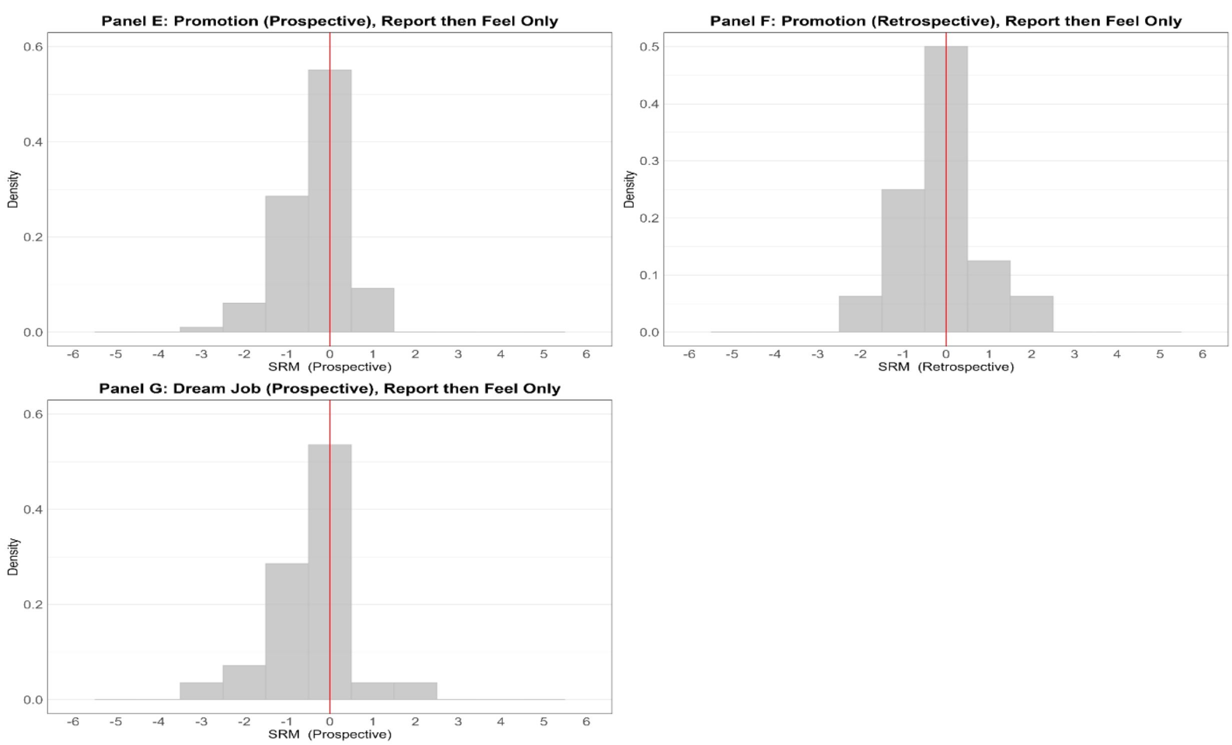
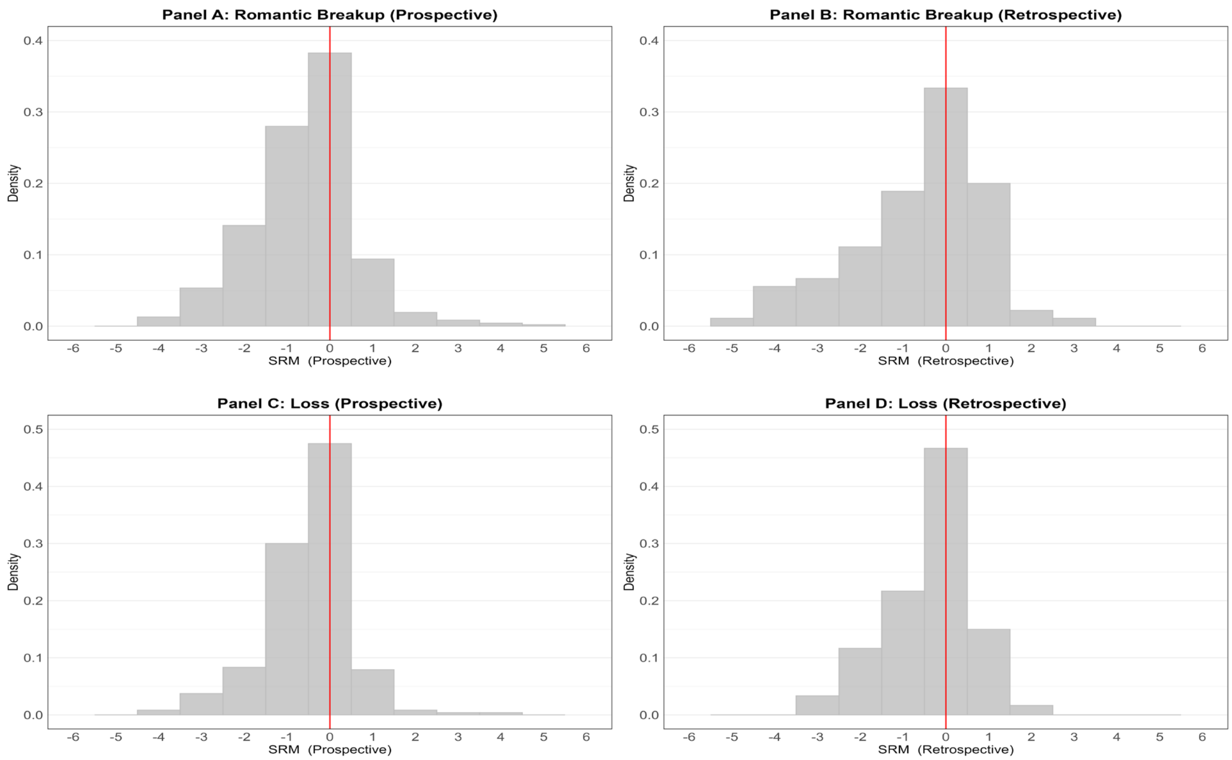


Figure A.4: Distributions for Self-Reported Misreporting (within subject), Pooled

Negative Events: Panels A through D



Positive Events: Panels E through G