Perspectives on Psychological Science
5(4) 343–355

DOI: 10.1177/1745691610374586
http://pps.sagepub.com

General Evaluability Theory

Christopher K. Hsee[1] and Jiao Zhang[2]
[1]Booth School of Business, University of Chicago, Chicago, IL, and [2]School of Business Administration, University of Miami, Miami, FL

Abstract

A central question in psychology and economics is the determination of whether individuals react differently to different values of a cared-about attribute (e.g., different income levels, different gas prices, and different ambient temperatures). Building on and significantly extending our earlier work on preference reversals between joint and separate evaluations, we propose a general evaluability theory (GET) that specifies when people are value sensitive and when people mispredict their own or others' value sensitivity. The GET can explain and unify many seemingly unrelated findings, ranging from duration neglect to affective forecasting errors and can generate many new research directions on topics ranging from temporal discounting to subjective well-being.

Keywords

evaluation mode, scope sensitivity, duration neglect, preference reversal, temporal discounting

Will people be more willing to take a vacation if the airfare is $200 rather than $300? Will people have a more painful memory of a surgery they underwent if the surgery lasted 2 hr rather than 1 hr? Will people work harder to earn $100 now or $120 a month from now? Will increasing the wealth of all improve the happiness of all?

These questions tap a diverse array of topics, ranging from price and demand, through duration neglect and temporal discounting, to wealth and happiness. Yet they all concern the same fundamental issue: value sensitivity.

The topic of value sensitivity is essential to both psychology and economics. Psychophysics (concerning how one reacts to sensory input), prospect theory (concerning how one values outcome and weighs probability), theory about price elasticity of demand (concerning how demand changes as price changes), and happiness research (concerning how happiness is related to income and other variables) are all about value sensitivity.

This article presents an integrative theory about value sensitivity—namely, about when one's evaluation is responsive to variations in value and when it is not. Value is a level of an objective attribute one cares about, such as price, duration, probability, wealth, and so on, and evaluation is one's subjective reaction, including willingness to purchase, willingness to work, happiness, and so on. If a person is willing to pay as much for a 1-karat diamond as he would for a 0.5-karat diamond, we say that the person is insensitive to diamond size (value) in that range. The more he is willing to pay for the larger diamond relative to the smaller diamond, the more he is sensitive to diamond size.

In what follows, we first introduce the theory, then show how it unifies many disparate findings in the existing literature, and finally suggest new research directions derived from the theory.

The Theory

Overview

Our theory, which we call the general evaluability theory (GET), seeks to address three broad questions: What factors determine value sensitivity? What is the relationship between these factors? Can people accurately predict value sensitivity? The GET starts from three basic propositions, each addressing one of the questions, and generates an array of specific hypotheses. The chart in Figure 1 offers an overview.

We should note here that GET greatly extends previous work on evaluability (e.g., Hsee 1996, 1998; Hsee, Blount, Loewenstein, & Bazerman, 1999). First, the notion of evaluability is now greatly expanded to three different types. Second, although the previous research on evaluability focused mainly on preference reversals, the present theory encompasses a much wider range of

Corresponding Author:
Christopher K. Hsee, Booth School of Business, University of Chicago, 5807 S. Woodlawn Avenue, Chicago, IL 60637
E-mail: chris.hsee@chicagobooth.edu

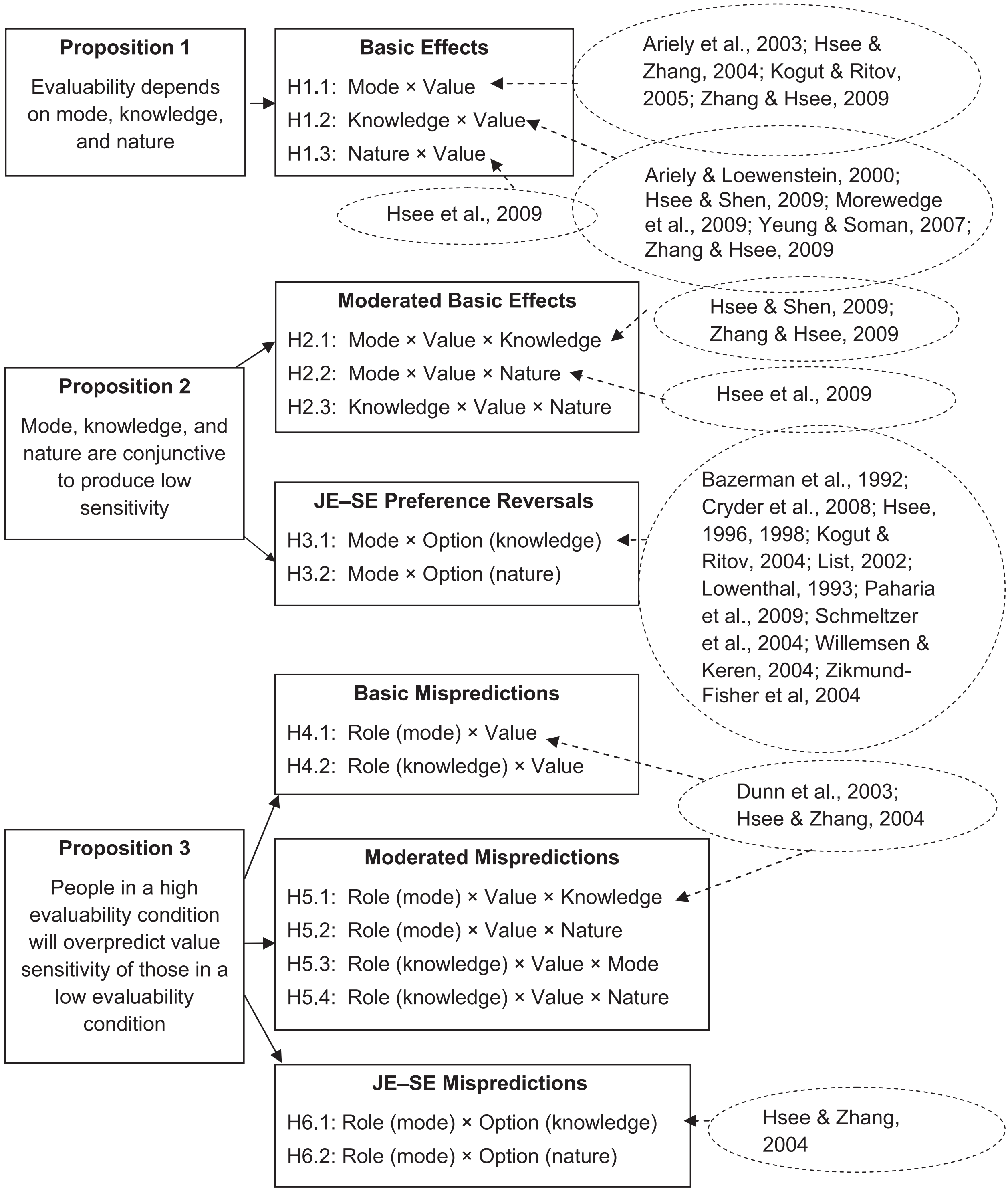

Fig. 1. Propositions and hypotheses of the general evaluability theory. JE = joint evaluation; SE = single evaluation.

topics, rendering preference reversals only an element of the current framework. Third, the current theory integrates numerous new findings that have emerged since the publications of the original papers on evaluability in the late 1990s. Finally, the current theory generates several new research directions which we believe are general and significant.

Proposition 1 (Evaluability)

We now introduce the three basic propositions of the GET and their corresponding hypotheses. The central tenet of our theory is that value sensitivity depends on evaluability. Evaluability refers to the extent to which a person has relevant reference

Table 1. Definitions of Evaluability Levels for Each of the Three Factors

Factor	Low-evaluability level	High-evaluability level
Mode	Single evaluation (SE)	Joint evaluation (JE)
Knowledge	Poor	Rich
Nature	Inherently inevaluable	Inherently evaluable

information to gauge the desirability of target values and map them onto evaluation. For instance, college students have abundant information about grade point average (GPA), such as its mean and range, but little knowledge about body-mass index (BMI, a measurement of height-adjusted body weight). Thus, for college students, evaluability of GPA is high and evaluability of BMI is low.[1]

Our first proposition concerns the determinants of evaluability:

> Proposition 1: Evaluability of, and hence value sensitivity to, values depends on three factors: mode, knowledge, and nature.

Each of the three factors can assume either a low- or high-evaluability level, as summarized in Table 1. For each factor, the low-evaluability level produces low-sensitivity evaluations and the high-evaluability level produces high-sensitivity evaluations, thus yielding a Factor × Value interaction (see the left panel of Figure 2—the figure is only illustrative; we do not assume that the slope for the low-evaluability level is entirely flat or that there is no main effect between the high- and the low-evaluability levels.) We now elaborate on each factor in turn.

Mode (H1.1). Any evaluation is formed in one of two modes or some combination of the two: the joint evaluation (JE) mode, in which two or more values are juxtaposed and compared, or the single evaluation (SE) mode, in which only a single value is presented and evaluated. Because evaluators can use one value as a reference to evaluate another value in JE, but not in SE, evaluability, and hence value sensitivity, will be higher in JE than in SE (H1.1). For example, individuals will be more sensitive to the size of a diamond in JE than they would in SE.

Knowledge (H1.2). Knowledge refers to the amount of distributional information (e.g., the range, the mean) about the target attribute one has acquired through learning. Knowledge may vary across individuals for the same attribute or across attributes for the same individual. Diamond traders know more about the sizes of diamonds than do ordinary consumers— an example of knowledge varying across individuals for the same attribute. College students know more about GPA than about BMI—an example of knowledge varying across attributes for the same individuals.

Evaluators have more reference information about an attribute if they have rich knowledge about the attribute. Thus, richer knowledge engenders higher evaluability and, hence, higher value sensitivity (H1.2). For example, everything else being equal, diamond traders (with rich knowledge) will have a stronger reaction toward a 0.5-karat diamond versus a 1-karat diamond than would regular consumers (with poor knowledge).

A general corollary of H1.2 is that, everything else being equal, evaluability is higher for values that constitute categorical differences rather than incremental differences. Two values are said to constitute a categorical difference if they differ in sign or in type, such as whether a stock price rises or falls, or whether an apartment features indoor parking or not. Two values are said to constitute an incremental difference if they differ only in degree, such as how much the stock price rises today, or how large the apartment is. Generally speaking, people are more knowledgeable about sign and type than about degree—for example, most people know that a positive investment return is good and a negative return is bad, but would not know as clearly how good a given positive return is or how bad a given negative return is. Thus, categorical differences lead to greater value sensitivity than do incremental differences. More precisely, people in SE are more sensitive to categorical differences than to incremental differences, holding their sensitivity to the two types of differences in JE constant. (See our discussion of the Mode × Value × Knowledge interaction later in this article.)[2]

Nature (H1.3). Nature refers to whether human beings have an innate and stable physiological or psychological "scale" (reference system) to evaluate values on an attribute. The attribute is inherently evaluable if they do or inherently inevaluable if they do not. Ambient temperature is an example of an inherently evaluable attribute; even without learning or social comparison, we can tell what temperature makes us comfortable and happy and what does not. Other examples include amount of sleep, social isolation, or connectedness. The size of a diamond and the power of a car are examples of inherently inevaluable attributes; without learning or comparison, we would not know how to assess such variables. Of course, some people know how to evaluate diamond size and car power, but such knowledge is learned, not innate. Because people possess innate reference systems for inherently evaluable attributes but not for inherently inevaluable attributes, value sensitivity (without learning or comparison) is higher for inherently evaluable attributes (H1.3). More precisely, people in SE are more sensitive to differences on an inherently evaluable attribute than to differences on an inherently inevaluable attribute, holding their sensitivity to the two types of differences in JE constant; see our discussion of the Mode × Value × Nature interaction later in this article.

It should be noted that classifying a variable as inherently evaluable does not mean that it is immune to the influence of external reference information (such as social comparison); instead, it means that people can evaluate the variable even without such information. Also, inherently evaluable variables are not always associated with basic biological needs—they also include socio-psychological variables, such as loneliness, depression, and sense of achievement. (For details, see Hsee, Yang, Li, & Shen, 2009.)

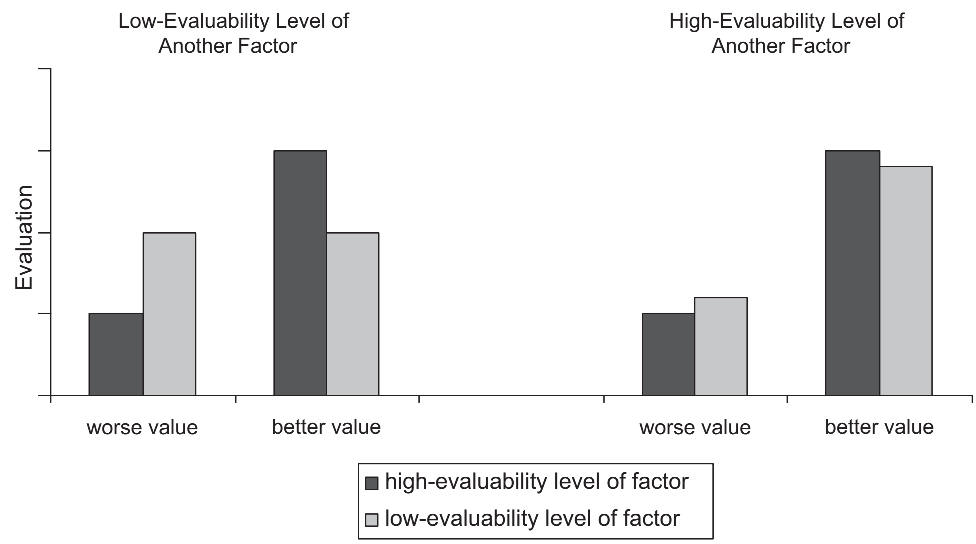

Fig. 2. Evaluability levels between factors. The left panel shows a basic Factor × Value interaction, where factor (the primary factor) can be mode, knowledge, or nature. The right panel shows that the basic two-way interaction can be moderated by another factor (a moderator), where the moderator can be any of the three factors other than the one in the basic two-way interaction.

All about reference information. Mode, knowledge, and nature are three sources of reference information. Mode is the most "here and now" source, concerning reference information that resides in the set of values or options presented; knowledge is less here and now, concerning reference information acquired through past experience; and nature is the least here and now source, concerning reference information that is innate, hard-wired, and rooted in evolution. Put differently, the three factors represent three types of evaluability: Mode is ad hoc evaluability, knowledge is learned evaluability, and nature is inherent evaluability. The distinction between learned evaluability (knowledge) and inherent evaluability (nature) is not an artificial classification, and it has important implications, which we will explore later.

Mode, knowledge, and nature are three independent dimensions and none of them can subsume another. Just as any physical object has three dimensions, length, height, and width, any evaluation assumes a value on each of the three factors. The evaluation is of either inherently evaluable or inherently inevaluable attributes, by evaluators with rich or poor knowledge, and in JE or SE.

Proposition 2: Relationship Among the Three Factors

According to our theory, value sensitivity will be high as long as the evaluator has the relevant reference information, regardless of whether the reference information comes from mode, knowledge, or nature. Thus, we propose the following:

> Proposition 2 (relationship among the factors): Mode, knowledge, and nature are conjunctive in producing low sensitivity. That is, the final evaluation will be of low sensitivity only if all three factors are of low evaluability.

Although the conditions producing low sensitivity appear stringent (requiring SE, poor knowledge, and inherent inevaluability), they are actually quite common: We often experience events one at a time, our knowledge is limited, and many variables we pursue accord no innate evaluation scales. For example, most of us live in only one home, know little about how to translate size into desirability, and have no innate scales to evaluate the size.

Proposition 2 generates two sets of hypotheses: one about moderators for the three basic two-way interactions and one about JE–SE preference reversals involving multiattribute options. We discuss these hypotheses in turn.

Moderated effects (H2.1–H2.3). Because the three factors are conjunctive in producing low value sensitivity, any one of the three basic Factor × Value interactions (H1.1–H1.3) can be moderated by another factor, as depicted in Figure 2. This yields the 3 three-way interactions described in H2.1–H2.3.

The Model × Value × Knowledge interaction predicts, for example, that individuals are more sensitive to diamond size in JE than in SE, but this effect is more pronounced if these individuals are diamond novices rather than diamond experts. It also predicts that people are more sensitive to the size of an apartment (an incremental difference) in JE than in SE, but their sensitivity to whether the apartment has indoor parking (a categorical difference) is relatively similar in JE and SE. Put differently, holding sensitivity to size and parking constant in JE, people in SE are more sensitive to parking than to size.

The Mode × Value × Nature interaction predicts that people are more sensitive to diamond size in JE than in SE, but they will be relatively equally sensitive to temperature in JE and in SE. Finally, the Knowledge × Value × Nature interaction proposes that learned knowledge about jewelry can greatly increase one's sensitivity to diamond size, but learned knowledge about temperature will not exert such a big effect on one's sensitivity to temperature, because organisms are already sensitive to temperature without external knowledge.

JE–SE preference reversals (H3.1–H3.2). All the hypotheses introduced so far concern evaluations of values on a single attribute. What about evaluations of two options that involve a tradeoff along two attributes? According to Proposition 2, preferences for the two options may reverse between JE and SE if one attribute is already of high evaluability in SE and there is little room for JE to bolster its evaluability, whereas the other attribute is of low evaluability in SE but JE considerably bolsters its evaluability.

There are two types of such JE–SE preference reversals: one involving knowledge (H3.1) and one involving nature (H3.2). The first constitutes a Mode (JE vs. SE) × Option (poor knowledge vs. rich knowledge attribute tradeoff) interaction, and the second is a Mode (JE vs. SE) × Option (inherently evaluable vs. inherently inevaluable attribute tradeoff) interaction.

Consider an early demonstration of the first type: The study (Hsee, 1996) involved evaluations of two job candidates for a computer programmer position, one with a higher GPA and the other with more programming experience. In SE, the candidate with the higher GPA was favored, whereas in JE, the candidate with more experience was favored. Presumably, the participants (all college students) possessed richer knowledge about GPA than about programming experience.

Proposition 3: Mispredictions

This proposition concerns prediction of one's own or another's value sensitivity. Technically, it is independent of the first two propositions, but incorporating it enriches our theory and allows us to explain more existing findings and generate more meaningful hypotheses.

Ample evidence indicates that if predictors and predictees (who can be either the predictors themselves at a future time or other individuals) are in different visceral states, predictors will project their own feelings onto the predictees and mispredict the predictees' feelings (e.g., Gilbert, Gill, & Wilson, 2002; Loewenstein, O'Donoghue, & Rabin, 2003; Van Boven, Loewenstein, & Dunning, 2005). For example, hungry diners tend to overestimate how much they will enjoy a rich dessert after a rich dinner even though they (should) know that they will be full after the dinner.

By the same token, if predictors and predictees are in different evaluability conditions, misprediction will also arise. Specifically,

> Proposition 3: If predictors are in a high-evaluability condition and predictees (either others or predictors themselves in the future) are in a low-evaluability condition, the predictors will overestimate the predictees' value sensitivity, even if the predictors know that the predictees are in a low-evaluability condition.

Note that the reverse is not applicable, namely, predictors in a low-evaluability condition (e.g., SE) will not mispredict the value sensitivity of predictees in a high-evaluability condition (e.g., JE). The reason is that to make a prediction for predictees in a high-evaluability condition, predictors will have to be informed of the additional reference information (e.g., the alternative option in JE) in the high-evaluability condition, and once they have this additional information, they are effectively in the same condition as the predictees.

Recall that Propositions 1 and 2 yield three sets of hypotheses: basic effects, moderated effects, and JE–SE preference reversals. Likewise, Proposition 3 yields three sets of hypotheses: basic mispredictions, moderated mispredictions, and JE–SE mispredictions.

Basic mispredictions (H4.1–H4.2). Theoretically, misprediction may happen if predictors are in the high-evaluability level and predictees are in the low-evaluability level on any of the three factors. However, no mispredictions can be made about nature because nature concerns different types of attributes. Thus, Proposition 3 yields only two basic mispredictions: one about mode and one about knowledge. The first posits that predictors in JE will overpredict the value sensitivity of predictees in SE, which implies a Role (mode) × Value interaction (H4.1). The second posits that predictors with rich knowledge will overpredict the value sensitivity of predictees with poor knowledge, which implies a Role (knowledge) × Value interaction (H4.2). Note that in each case, predictors always assume the high-evaluability level of the factor and predictees always assume the low-evaluability level.

Moderated mispredictions (H5.1–H5.4). In the section detailing Proposition 2, we explained that any of the basic Factor × Value interactions could be moderated by another factor. Likewise, either basic misprediction can be moderated by another factor, yielding four possible combinations, as described in H5.1–H5.4.[3] Each of these hypotheses is ecologically meaningful and testable—take H5.1 for example. Suppose that a man is in a jewelry store shopping for a diamond ring for his fiancée and is comparing diamonds of different sizes. Chances are that he will overpredict his fiancée's sensitivity to the size of the diamond and spend too much on a big diamond because he is in JE when buying the diamond whereas she is in SE upon receiving it (H4.1). However, this misprediction will be less pronounced if they (especially she) are knowledgeable about diamonds (e.g., she was married and had received diamond rings before) than it would be if she is not knowledgeable.

JE–SE mispredictions (H6.1–H6.2). Just as JE–SE preference reversals may arise if two options vary along a high-evaluability attribute and a low-evaluability attribute, so will JE–SE mispredictions occur for such options if the predictor is in JE and the predictee in SE. Specifically, there are two possible JE–SE mispredictions, one involving knowledge (H6.1) and one involving nature (H6.2). To illustrate, H6.2 posits that if two options involve a trade-off between an inherently-evaluable attribute and an inherently-inevaluable attribute, and if predictors are in JE while predictees are in SE, then the predictors may mistakenly predict the predictees to favor the option superior on the inherently inevaluable attribute.

Integrating Existing Findings

The GET can parsimoniously explain various seemingly unrelated findings. Below are some examples.

Scope Insensitivity and Duration Neglect

People are sometimes clueless about the "scope" (e.g., frequency, size, or duration) of an event or outcome that they care about. People are willing to pay as much to save 2,000 birds as they would to save 200,000 birds (Desvousges, Johnson, Dunford, Boyle, & Wilson, 1993). They are willing to donate as much to help a single victim as they would to help a group of victims (Kogut & Ritov, 2005). However, these dramatic scope-insensitivity effects occur only when different values (e.g., 2,000 birds vs. 200,000 birds) are presented in SE, and they disappear when they are presented in JE. In our opinion, these results are manifestations of the Mode $\times$ Value interaction in H1.1.

The most studied form of scope insensitivity is *duration neglect*, which states that retrospective evaluation of an event is insensitive to its duration (e.g., Fredrickson & Kahneman, 1993; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). In a well-known experiment (Kahneman et al., 1993), respondents experienced two episodes one at a time (in SE though within-participants). (By default, temporally extended events are experienced one at a time, as in SE or quasi-SE.) One episode (the shorter one) required participants to immerse their hands in very cold water for 60 s; the other episode (the longer one) required them to immerse their hands in very cold water for 60 s and then in mildly cold water for another 30 s. Objectively the longer episode was worse, but respondents remembered it as being less aversive and were more willing to repeat it, indicating sensitivity to the final intensity (very cold or mildly cold) of the experience but insensitivity to the duration (60 s or 90 s).

We suggest that a possible reason for duration neglect is that duration is inherently inevaluable and it is also low in learned evaluability—in most experiments showing duration neglect, the stimuli are unfamiliar to the participants. Recently, Morewedge, Kassam, Hsee, and Caruso (2009) asked respondents to retrospectively evaluate a noise, which was either unlabeled or labeled as an Australian phone ring, that lasted either 7 s or 21 s. They replicated the typical duration neglect finding in the no-label condition but observed significant duration sensitivity in the labeled condition. Presumably the labeling prompted respondents to use the duration of a usual (familiar) phone ring as a reference, turning a poor-knowledge experience into a rich-knowledge experience. This result reveals the Knowledge $\times$ Value interaction in H1.2. Similarly, Ariely and Loewenstein (2000) found that participants were sensitive to the durations of noises if they compared the noises with a referent noise.

More generally, we propose that the intensity of an event is inherently easier to evaluate than the duration of the event. Humans have an innate psychophysical scale to tell how loud a noise feels aversive and how strong an electrical shock is painful, but do not have such an inherent reference system to evaluate their durations. Most individuals would find a 100-dB noise aversive but would not have much idea whether a 100-s noise was long or short. That is probably why we often observe sensitivity to intensity and insensitivity to duration.

JE–SE Preference Reversals and Arbitrary Cohesiveness

Intuitively, two items (e.g., two job candidates) evaluated comparatively (i.e., in JE) or separately (i.e., in SE) should yield the same preference ranking. If Candidate A is rated more favorably than Candidate B in JE, then A should also be rated more favorably than B in SE. In reality, preference ranking often reverses between JE and SE (e.g., Bazerman, Loewenstein, & Blount White, 1992; Bazerman, Schroth, Shah, Diekmann, & Tenbrunsel, 1994; Bazerman, Tenbrunsel, & Wade-Benozi, 1998; Chatterjee, Heath, & Min, 2009; Cryder, Mullen, & Loewenstein, 2008; Gonzalez-Vallejo & Moran, 2001; Hsee, 1996, 1998; Hsee et al., 1999; Irwin, Slovic, Lichenstein, & McClelland, 1993; Kahneman & Ritov, 1994; Kogut & Ritov, 2005; List, 2002; Loewenthal, 1993; Moore, 1999; Paharia, Kassam, Greene, & Bazerman, 2009; Schkade & Kahneman, 1998; Schmeltzer, Caverni, & Warglien, 2004; Tehila & Ritov, 2005; Willemsen & Keren, 2004; Zikmund-Fisher, Fagerlin, & Ubel, 2004; see Larrick, 2004, for a discussion on how to minimize these and other biases).

For example, Zikmund-Fisher et al. (2004) reported that patients in SE favored less experienced surgeons who graduated from Harvard over more experienced surgeons who graduated from the University of Iowa, though in JE their preference were reversed. Presumably, patients had more knowledge about the quality of universities than about the experience of a doctor. This, like many other such findings, is a manifestation of the Mode $\times$ Option (knowledge) interaction in H3.1.

As mentioned earlier, a special case of the JE–SE preference reversals in H3.1 arises when the options to be evaluated involve an incremental difference on one attribute and a categorical difference on another. For example, List (2002) examined bidding prices for two sets of baseball cards, one containing 10 cards, all of which were described as being in mint condition, and the other containing 13 cards, of which 10 were described as being in mint condition and 3 were described as being in poor condition. In SE, bidding was higher for the pack of 10 mint-condition cards; in JE, bidding was higher for the pack of 13 mixed-condition cards. Notice that whether a set included 10 or 13 cards was merely an incremental difference, and whether the set contained all mint-condition cards or contained some poor-condition cards was a categorical difference. The baseball card study was a replication of a series of original studies by Hsee (1998) showing the same type of preference reversals. In one study, participants in SE were willing to pay more for a 24-piece intact dinnerware set than for a 40-piece dinnerware set that contained a few broken pieces, yet their preferences reversed in JE. In another study, participants in SE were willing to pay more for 7 oz of ice cream served in a 5-oz cup (overfilled) than for 8 oz of ice cream served in a 10-oz cup (underfilled), but their preference reversed in JE.

More such JE–SE reversals have been reported recently. For example, Paharia et al. (2009) reported that participants in SE judged companies that caused direct and less severe harm to consumers more negatively than companies that caused indirect (through an agent) but more severe harm, whereas in JE their judgment was reversed. Cryder et al. (2008) found that a hedonic but less expensive reward (e.g., $1.49 chocolate) was more effective in SE than a nonhedonic but more expensive reward (e.g., $3 cash) in motivating participants to complete a survey, whereas in JE the latter reward was more effective. The authors explained their findings in terms of a conflict between what participants wanted and what they thought they should choose, but these findings are also consistent with our hypothesis (H3.1). In Paharia et al.'s study, whether the harm was direct or through a third party was a categorical difference, whereas the difference in the severity of the harm was incremental. In Cryder et al.'s study, whether the reward was hedonic or not was categorical, whereas the difference in the worth of the reward was incremental. Indeed, most, if not all, of the JE–SE preference reversals documented in the existing literature can be parsimoniously explained by the GET.

Another general phenomenon that seems to be consistent with the GET is what Ariely, Loewenstein, and Prelec (2003) call "coherent arbitrariness." In an ingenious study, the authors first had respondents listen to a 30-s noise, then asked if they were willing to hear the same noise again for $.10 or $.50 (between-subjects manipulation), and finally asked how much they would demand to hear longer and shorter noises (within-subjects manipulation). Participants' demands were heavily influenced by the arbitrary anchor, demanding more in the $.50 condition than in the $.10 condition, yet were coherent with the durations of the noises—the participants demanded more money for longer noises. These findings are consistent with our hypothesis that people exhibit low value sensitivity in SE and high value sensitivity in JE. Notice that the anchor ($.10 or $.50) was manipulated between subjects (SE) and the durations were manipulated within subjects (JE). Thus, participants in SE were insensitive to how much they should be paid and so were influenced by the anchor but they were sensitive to duration variations in JE and therefore demanded more for longer noises.

Misprediction and Mischoice

Biases in affective forecasting (prediction of future hedonic experiences) have been the focus of a large literature in social psychology and decision theory (e.g., Gilbert et al., 2002; Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Kahneman & Snell, 1992; Loewenstein et al., 2003; Wilson & Gilbert, 2003).

A general type of forecasting error is the distinction bias: People faced with multiple options (i.e., in JE) overpredict the value sensitivity of their future selves or of others who experience only one of the options (i.e., in SE; Hsee & Zhang, 2004). For example, Dunn, Wilson, and Gilbert (2003) observed that before finding out which of 12 dormitory buildings they would be assigned to, college freshmen predicted that the physical features of the buildings would have a great impact on their experience when they lived in each building; however, after they had been living in the buildings assigned to them, the physical features had little impact. One possible explanation of this intriguing finding is that the freshmen were in JE when making the predictions and in SE when living in the assigned dorm and that they were more sensitive to physical variations in JE than in SE, which is a manifestation of the Role (mode) $\times$ Value interaction in H5.1. The misprediction in this study may also result from other factors, such as failure to predict hedonic adaption and failure to consider social-life differences between the buildings.

In a more controlled test of the distinction bias (Hsee & Zhang, 2004), participants were assigned to be either predictors or experiencers. The predictors were shown four lists of words, which varied both in valence (positive or negative words) and in length (10 or 25 words), and were asked to predict how a person would feel after reading only one of the lists. Each experiencer read only one of the lists and reported feelings. Notice that the predictors were in JE and the experiencers were in SE. The results matched the Role (mode) $\times$ Value $\times$ Knowledge interaction in H5.2: The predictors thought that SE experiencers would be sensitive to both the valence of the lists (a categorical difference) and the length of the lists (an incremental difference), yet SE experiencers were sensitive only to valence, not to length.

Mispredictions can breed mischoices. Choice is typically made in JE, whereas the consequence of the choice is often experienced in SE. For example, when buying a sofa, we usually compare multiple options, whereas when using the sofa we have eventually bought, we face that sofa alone. According to H6.1, when two options involve a tradeoff between a rich-knowledge attribute and a poor-knowledge attribute, predictors in JE are more likely to favor the option superior on the poor-knowledge attribute than will experiencers in SE. Similarly, according to H6.2, when two options involve a trade-off between an inherently evaluable attribute and an inherently inevaluable attribute, predictors in JE are more likely to favor the option superior on the inherently inevaluable attribute. To the extent that people resolve choices by relying on their predictions, mispredictions as described in H6.1 and H6.2 will result in mischoices.

Such a mischoice is demonstrated in Hsee and Zhang (2004). Participants in the study were asked to choose between two tasks: one asking them to recall a sad personal story and eat a large piece of chocolate and the other to recall a happy

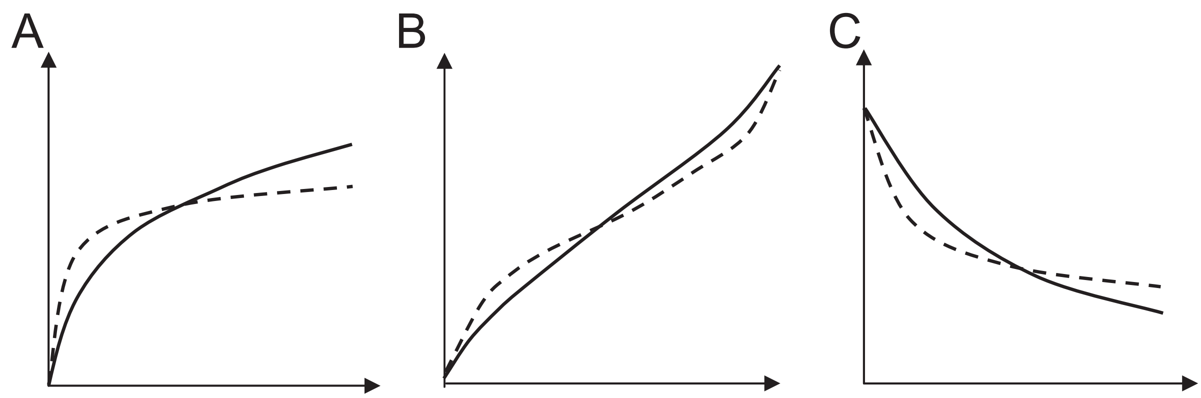

Fig. 3. Value (A), probability weighting (B), and temporal discounting (C) functions. For all three functions, the curve is more linear in joint evaluation (solid line) than in single evaluation (dotted line).

personal story and eat a small piece of chocolate. The result revealed a stark choice–experience inconsistency: When making the choice, most participants chose the sad-story/large-chocolate task, but when doing the chosen task, participants who chose the happy-story/small-chocolate task were happier than were those who chose the sad-story/large-chocolate task. This pattern is consistent with H6.1: The difference in the size of a piece of chocolate is incremental and the difference in the nature of the story (sad vs. happy) is categorical.

Relative and Absolute Determinants of Happiness

Over the past decades, scholars have vigorously argued about whether happiness is relative or absolute. Some (e.g., Easterlin, 1974, 1995) believe that happiness depends only on relative outcomes, such as social comparison (Festinger, 1954); for example, how happy people feel about their incomes depends only on how their incomes compare with their neighbors', their friends', and their colleagues'. Others (e.g., Veenhoven, 1991) believe that happiness depends on absolute outcomes and does not rely on social comparison. Which view is correct? A recent study (Hsee et al., 2009) suggests that the answer depends on the nature of the variable.

In the study, residents in 31 representative cities in China were approached during the winter in a large telephone survey and asked about their room temperature, the monetary worth of their jewelry, and their happiness with these variables. The authors analyzed the effects of temperature and jewelry value on happiness both within cities and across cities, observing that residents in China would compare with each other within a city but not across cities, and hence, within-city evaluations mimicked JE and between-city evaluations mimicked SE. The result for temperature emerged to be highly different from that for jewelry value. Temperature had both within-city and between-city effects, but jewelry value had only a within-city effect, suggesting that temperature could influence happiness even without social comparison (between cities) and jewelry influenced happiness only when social comparison was easy (within cities). These findings reflect the Mode × Value × Nature interaction in H2.2, where room temperature is inherently evaluable and jewelry value is inherently inevaluable.

The findings help answer the question of whether happiness depends only on relative consumption levels or also on absolute consumption levels: For inherently inevaluable attributes such as the monetary worth of possessed jewelries, only relative consumption levels influence happiness—that is, happiness depends only on social comparison. In contrast, for inherently evaluable attributes, such as temperature, both relative and absolute consumption levels influence happiness.

Directions for Future Research

To illustrate new research directions that the GET can generate, we discuss three overarching topics: valuation, temporal discounting and probability weighting; price and demand; and wealth and happiness.

Valuation, Temporal Discounting, and Probability Weighting

Arguably the three most important classes of functions in judgment and choice are value (utility) functions, temporal discounting functions, and probability weighting functions. Value functions describe how objective quantities map onto subjective evaluations (e.g., Kahneman & Tversky, 1979), temporal discounting functions describe how decision makers discount the values of temporally delayed outcomes (e.g., Fredrick, Loewenstein, & O'Donoghue, 2002; Read, 2004; Soman et al., 2005), and the probability weighting functions describe how probabilities translate into decision weights (e.g., Kahneman & Tversky, 1979).

These three types of functions share a common characteristic: They are all nonlinear, especially around endpoints—zero quantity in value, no delay in temporal discounting, and impossibility (0%) or certainty (100%) in probability weighting (Fig. 3). A large body of research has examined factors that influence the nonlinearity of each of the three types of functions (e.g., Gonzalez & Wu, 1999; Hsee & Rottenstreich, 2004; Kluger, Stephan, Ganzach, & Hershkovitz, 2004; Read, Fredrick, Orsel, & Rahman, 2005; Rottenstreich & Hsee, 2001).

We now propose another factor that influences the nonlinearity of all of the three types of functions: evaluation mode. All of these functions are less linear in SE than in JE.

Value functions. We propose that value functions are more linear in JE than in SE, as depicted in the left panel of Figure 3. In JE, people will be happier with $10 than with $0 or with $20 than with $10, and the two differences in happiness will be similar. In SE, people will still be happier with $10 than with $0, but will be similarly happy with $20 and with $10. The difference between nothing ($0) and something (e.g., $10) is categorical, whereas the difference between something small and something larger (e.g., $10 vs. $20) is incremental. According to H2.1, people in JE will be similarly sensitive to the incremental difference and to the categorical difference, and people in SE will be sensitive to the categorical difference but insensitive to the incremental difference. We have evidence for this prediction (Hsee & Zhang, 2004; Zhang & Hsee, 2009). In a recent study (Zhang & Hsee, 2009), participants in JE evaluated payoffs of different magnitudes (0 chocolates, 5 chocolates, and 10 chocolates) rather linearly, whereas those in SE evaluated 0 chocolates and 5 chocolates quite differently but evaluated 5 chocolates and 10 chocolates similarly.

Probability weighting functions. Just as value functions are less linear in SE than in JE, we propose that probability weighting functions are also less linear in SE than in JE, as depicted in the middle panel of Figure 3. The reason is similar: Impossibility (0%) versus some possibility is a categorical difference, as is some possibility versus certainty (100%), but a small possibility versus a larger possibility is only an incremental difference. According to H2.1, we predict that the probability weighting function will be more S-shaped in SE than in JE. For example, in JE, one will be happier with $100 guaranteed than with a 90% chance at $100, and happier with a 90% chance at $100 than with an 80% chance at $100. In SE, however, one will be happier with $100 guaranteed than with a 90% chance at $100, but will feel similarly happy with a 90% chance at $100 and with an 80% chance at $100.

Temporal discounting functions. As with the value and the probability weighting functions, we propose that the temporal discounting functions are also less linear in SE than in JE, as shown in the right panel of Figure 3. The difference between no delay and some delay is categorical, whereas the difference between a short delay and a longer delay is merely incremental. According to H2.1, we predict that around the zero (no delay) point, the temporal discounting function will be more convex in SE than in JE. For example, in JE one will feel happier with $100 now than with $100 in 2 weeks, and happier with $100 in 2 weeks than with $100 in 4 weeks. In SE, however, one will feel happier with $100 now than with $100 in 2 weeks, but will feel similarly happy with $100 in 2 weeks and with $100 in 4 weeks.

From the analysis above, we can also expect intriguing JE–SE preference reversals for options involving tradeoffs between value, delay, and/or probability. For example, consider two options: A = receiving $100 guaranteed and B = a 90% chance at receiving $150. On the basis of H3.1, we expect A to be favored in SE and B to be favored in JE, because the difference in probability (100% vs. 90%) is categorical and the difference in payoff ($100 vs. $150) is incremental. Similarly, consider another pair of options: C = receiving $200 now, and D = receiving $250 in 3 weeks. We expect C to be favored in SE and D to be favored in JE, because the difference in delay (no delay vs. a 3-week delay) is categorical and the difference in payoff ($200 vs. $250) is incremental.

To the best of our knowledge, we are the first to suggest that mode influences the shapes of all three types of functions. Future research testing the predictions delineated in this section may yield fruitful results.

Price and Demand

The relationship between price and demand is a fundamental issue in both microeconomics and marketing. It is generally believed that the higher the price of a good, the lower the demand for it, holding supply constant. The sensitivity of demand to price is called price elasticity. Economists and marketers have identified various determinants of price elasticity, such as necessity of the good and the wealth of the consumer. Besides these factors, our theory suggests another determinant of price elasticity: price evaluability.

The price evaluability of a good depends on whether consumers are knowledgeable about the price of that good. For example, most city dwellers are more knowledgeable about taxi fares than about horse-ride fares. Thus, according to our theory (H1.2), price elasticity will be higher for taxi rides than for horse rides. The finding of a recent study (Study 1 in Hsee & Shen, 2009) is consistent with this prediction. In the study, participants were either asked to indicate whether they would take a taxi or a horse ride to a destination. The price for the ride, low or high, was manipulated between participants. A pretest confirmed that participants were knowledgeable about taxi fares but unknowledgeable about horse-ride fares. As predicted, willingness to take a taxi was highly sensitive to the price of the taxi, whereas willingness to take a horse ride was insensitive to the price of the horse ride.

Can marketers and sellers accurately predict buyers' price elasticity? There is reason to believe that marketers and sellers are likely to overpredict buyers' price elasticity. When setting the price for a good, sellers typically entertain different prices before settling on a price. Potential buyers of the good, on the other hand, only see the price the sellers have set. In other words, the sellers are in JE and the buyers are in SE. According to our theory (H6.1), sellers will overpredict buyers' price elasticity if buyers have poor knowledge about the price, but not if buyers have rich knowledge about the price. Overpredicting buyers' price sensitivities can lead sellers to underprice their goods and consequently forgo profits. Thus, the managerial implication is that sellers should set higher prices than they typically do for unfamiliar products (e.g., horse rides, a new type of cell phone that transmits smell), but not for familiar products (e.g., taxi, a conventional cell phone).

Wealth and Happiness

A question that weighs on the minds of many is whether increasing the wealth from one generation to the next makes the next generation happier. Numerous scholars have conducted important research on this question (e.g., Diener & Biswas-Diener, 2002; Diener, Kahneman, Tov, & Arora, in press; Diener, Ng, Harter, & Arora, in press; Diener, Sandvik, Seidlitz, & Diener, 1993; Easterlin, 1974, 1995; Veenhoven, 1991) and have produced mixed findings. Some studies show that increasing the wealth of all does not increase happiness of all (e.g., Easterlin, 1974, 1995). Other studies suggest that raising wealth can increase reported subjective well-being and that life satisfaction in developed countries is on average higher than that in less developed ones (e.g., Clark, Frijters, & Shields, 2008; Diener, Kahneman, et al., in press; Diener, Ng, et al., in press; Diener et al., 1993; Kahneman, 2008; Stevenson & Wolfers, 2008).

According to our theory, a cross-generation increase in wealth may or may not lead to a cross-generation increase in happiness depending on how the increased wealth is used. If it is used to improve inherently evaluable aspects of life, such as sanitary systems, indoor temperature, childhood mortality rate, and physiological and psychological health, it will increase happiness across time; otherwise, it may not. The finding of the temperature/jewelry study reviewed earlier (Hsee et al., 2009) is consistent with this proposition. In that study, each city can be compared to a separate generation, with richer cities to richer generations. The finding suggests that although both room temperature and jewelry value exert significant within-generation effects, only room temperature exerts a between-generation effect.

One might wonder why we only recommend improving inherently evaluable aspects of life and not learned-evaluable aspects of life. The reason is that inherent evaluability (nature) and learned evaluability (knowledge) differ in a critical way. Individuals' reference systems and evaluation scales for inherently evaluable variables (e.g., what temperature feels comfortable) are relatively stable across generations, whereas their reference system and evaluation scale for learned-evaluable variables are time-specific and will change as social norms about these variables (e.g., which diamond size is considered desirable) change. This distinction has not been made in the previous research on evaluability (e.g., Hsee, 1996; Hsee et al., 1999; Yeung & Soman, 2005, 2007). Consequently, improving inherently evaluable variables from one generation to the next will make the next generation absolutely happier, and improving learned-evaluable variables is merely a zero-sum game and will only produce a treadmill effect.

Recently, Diener, Kahneman, et al. (in press) and Diener, Ng, et al. (in press) found positive correlations between income and happiness both across countries and across generations in the last few decades. How can our theory account for these positive correlations? One viable explanation, as mentioned earlier, is that higher income has led to improvement of inherently evaluable variables, allowing individuals living in richer regions or richer times to enjoy better inherently evaluable aspects of life, and because evaluation scales for these variables are not region- or time-specific, these individuals are absolutely happier than those living in poorer regions or times. Another possible reason is that people in different countries may share some common standards of material well-being, such as having a television, having a car, and having one's own apartment in a city, and the standards, though socially determined, may have remained relatively stable over the last few decades. Thus, as the percentage of people owning these material goods increases, the average happiness also increases. However, we suspect that the standards for such learned-evaluable variables will sooner or later change as time evolves, whereas standards for inherently evaluable variables will not.

One may also wonder if the gain in happiness brought about by improving inherently evaluable aspects of life will be eventually erased by hedonic adaptation (e.g., Brickman & Campbell, 1971; Brickman, Coates, & Janoff-Bulman, 1978; Diener, Lucas, & Scollon, 2006; Helson, 1964). We doubt it. Consider the findings of the temperature/jewelry study. Respondents in that study had presumably experienced their room temperature many times and adapted to it. However, the respondents still showed a significant between-city effect of room temperature on their happiness.

One of the major puzzles in happiness research is determining which variables are resistant to hedonic adaptation and which are prone to it. Our research can potentially shed light on this question. We conjecture that the rate of hedonic adaptation is related to the degree of inherent evaluability. Specifically, inherently evaluable attributes are more resistant to hedonic adaptation than inherently inevaluable attributes, because the former are associated with stable innate evaluation scales and the latter are not. Inherently evaluable attributes are like a foam ball. The foam ball has an inherent shape; although the ball can be temporally squeezed by an external force, it will eventually return to its original shape. This analysis implies that inherent evaluability is a cause, not a result of (lack of) hedonic adaptation. That is, inherent evaluability can explain why some variables are more resistant to hedonic adaption than others, yet hedonic adaptation cannot explain why some variables are more inherently evaluable than others.

Summary

Many diverse and important problems such as the relationship between demand and price, between wealth and happiness, and between time duration and retrospective evaluation concern the same fundamental question: When are people value sensitive? The GET advances our understanding of this fundamental question. According to the theory, evaluability, hence value sensitivity, is high if evaluators have acquired relevant reference information through any of three main sources of reference information—mode, knowledge, and nature. Furthermore, systematic overprediction of value sensitivity will arise if predictors are in a high-evaluability condition and predictees are in a low-evaluability condition.

In our opinion, the GET is parsimonious and generative. It is parsimonious because it integrates several disparate lines of research with only three basic factors. And it is generative because the interplay between these factors yields a wide array of testable hypotheses and provides a fertile ground for future research.

Notes

1. We assume in this article that evaluators at least know what the variable is about and which direction is more desirable.
2. In Hsee and Zhang (2004), categorical differences are referred to as qualitative differences, and incremental differences are referred to as quantitative differences.
3. There are four moderated mispredictions here, but only three moderated effects in Proposition 2 because the order of factors (determining which is the basic factor and which is the moderator) does not matter for moderated effects but does matter for moderated mispredictions.

Acknowledgments

The authors thank the University of Chicago Center for Decision Research and the Templeton Foundation for research support, and thank Eva Buechel, Zach Burns, Paulina Celedon, Dan Gilbert, and Lucy Shen for helpful suggestions on earlier drafts.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

Ariely, D., & Loewenstein, G. (2000). When does duration matter in judgment and decision making? *Journal of Experimental Psychology: General*, *129*, 508–529.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *Quarterly Journal of Economics*, *118*, 73–105.

Bazerman, M.H., Loewenstein, G., & Blount White, S. (1992). Reversals of preference in allocation decisions: Judging an alternative versus choosing among alternative. *Administrative Science Quarterly*, *37*, 220–240.

Bazerman, M.H., Schroth, H.A., Shah, P.P., Diekmann, K.A., & Tenbrunsel, A.E. (1994). The inconsistent role of comparison others and procedural job descriptions: Implications for job acceptance decisions. *Organizational Behavior and Human Decision Processes*, *60*, 326–352.

Bazerman, M.H., Tenbrunsel, A.E., & Wade-Benozi, K.A. (1998). Negotiation with yourself and losing: Understanding and managing competing internal preferences. *Academic Management Review*, *23*, 225–241.

Brickman, P., & Campbell, D.T. (1971). Hedonic relativism and planning the good society. In M.H. Appley (Ed.), *Adaptation level theory: A symposium* (pp. 287–302). New York: Academic Press.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, *36*, 917–927.

Chatterjee, S., Heath, T.B., & Min, J. (2009). The susceptibility of mental accounting principles to evaluation mode effects. *Journal of Behavioral Decision Making*, *22*, 120–137.

Clark, A.E., Frijters, P., & Shields, M. (2008). Relative income, happiness and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, *46*, 95–144.

Cryder, C.E., Mullen, E.E., & Loewenstein, G. (2008, November). *Wanting versus choosing: A disconnect between what moves us and what we prefer*. Paper presented at the Society for Judgment and Decision Making Preconference: Using Human Nature to Improve Human Life, Chicago, IL.

Desvousges, W.H., Johnson, F.R., Dunford, R.W., Boyle, S.P., & Wilson, K.N. (1993). Measuring resource damages with contingent valuation. Tests of validity and reliability. In J.A. Hausman (Ed.), *Contingent valuation: A critical assessment* (pp. 91–164). Amsterdam: North-Holland.

Diener, E., & Biswas-Diener, R. (2002). Will money increase subjective well-being? *Social Indicators Research*, *57*, 119–169.

Diener, E., Kahneman, D., Tov, W., & Arora, R. (in press). Income's association with judgments of life versus feelings. In E. Diener, J. Helliwell, & D. Kahneman (Eds.), *International differences in well-being* New York: Oxford University Press.

Diener, E., Lucas, R.E., & Scollon, C.N. (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. *American Psychologist*, *61*, 305–314.

Diener, E., Ng, W., Harter, J., & Arora, R. (in press). Wealth and happiness across the world: Material prosperity predicts life evaluation, while psychosocial prosperity predicts positive feeling. *Journal of Personality and Social Psychology*.

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute? *Social Indicators Research*, *28*, 195–223.

Dunn, E.W., Wilson, T.D., & Gilbert, D.T. (2003). Location, location, location: The misprediction of satisfaction in housing lotteries. *Personality and Social Psychology Bulletin*, *29*, 1421–1432.

Easterlin, R.A. (1974). Does economic growth improve the human lot? In P.A. David & M.W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz* (pp. 89–125). New York: Academic Press.

Easterlin, R.A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization*, *27*, 35–47.

Festinger, L. (1954). A theory of social comparison. *Human Relations*, *7*, 117–140.

Fredrick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*, 351–401.

Fredrickson, B.L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, *65*, 45–55.

Gilbert, D.T., Gill, M.J., & Wilson, T.D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes*, *88*, 430–444.

Gilbert, D.T., Pinel, E.C., Wilson, T.D., Blumberg, S.J., & Wheatley, T. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*, 617–638.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129–166.

Gonzalez-Vallejo, C., & Moran, E. (2001). The evaluability hypothesis revisited: Joint and separate evaluation preference reversal as a function of attribute importance. *Organizational Behavior and Human Decision Processes*, *86*, 216–233.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Hsee, C.K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, *67*, 247–257.

Hsee, C.K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, *11*, 107–121.

Hsee, C.K., Blount, S., Loewenstein, G., & Bazerman, M. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *125*, 576–590.

Hsee, C.K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General*, *133*, 23–30.

Hsee, C.K., & Shen, L. (2009). *Taxi versus horse: What affects consumer price sensitivity and when price-setters mispredict it?* Unpublished manuscript.

Hsee, C.K., Yang, Y., Li, N., & Shen, L. (2009). Wealth, warmth, and well-being: Whether happiness is relative or absolute depends on whether it is about money, acquisition, or consumption. *Journal of Marketing Research*, *46*, 396–409.

Hsee, C.K., & Zhang, J. (2004). Distinction bias: Misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology*, *86*, 680–695.

Irwin, J.R., Slovic, P., Lichenstein, S., & McClelland, G. (1993). Preference reversals and the measurement of environmental values. *Journal of Risk and Uncertainty*, *6*, 5–18.

Kahneman, D. (2008). *The sad tale of the aspiration treadmill*. Retrieved May 21, 2009, from http://www.edge.org/q2008/q08_17.html#kahneman

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

Kahneman, D., & Ritov, I. (1994). Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty*, *9*, 5–38.

Kahneman, D., & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making*, *5*, 187–200.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–292.

Kluger, A.N., Stephan, E., Ganzach, Y., & Hershkovitz, M. (2004). The effect of regulatory focus on the shape of probability-weighting function: Evidence from a cross-modality matching method. *Organizational Behavior and Human Decision Processes*, *95*, 20–39.

Kogut, T., & Ritov, I. (2005). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, *97*, 106–116.

Larrick, R.P. (2004). Debiasing. In: D.J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316–337). Oxford, United Kingdom: Blackwell.

List, J. (2002). Preference reversals of a different kind: The "more is less" phenomenon. *American Economic Review*, *92*, 1636–1643.

Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Projection bias in predicting future utility. *Quarterly Journal of Economics*, *118*, 1209–1248.

Loewenthal, D. (1993). *Reversals of preference in candidate choice*. Unpublished manuscript, Carnegie-Mellon University, Pittsburgh, PA.

Moore, D.A. (1999). Order effects in preference judgments: Evidence for context dependence in the generation of preferences. *Organizational Behavior and Human Decision Processes*, *78*, 146–165.

Morewedge, C.K., Kassam, K.S., Hsee, C.K., & Caruso, E.M. (2009). Duration sensitivity depends on stimulus familiarity. *Journal of Experimental Psychology: General*, *138*, 177–186.

Paharia, N., Kassam, K.S., Greene, J.D., & Bazerman, M.H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes*, *109*, 134–141.

Read, D. (2004). Intertemporal choice. In: D.J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 424–443). Oxford, United Kingdom: Blackwell.

Read, D., Frederick, S., Orsel, B., & Rahman, J. (2005). Four scores and seven years from now: The date/delay effect in temporal discounting. *Management Science*, *51*, 1326–1335.

Rottenstreich, Y., & Hsee, C.K. (2001). Money, kisses, and electronic shocks: On the affective psychology of risk. *Psychological Science*, *12*, 185–190.

Schkade, D., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, *9*, 340–346.

Schmeltzer, C., Caverni, J., & Warglien, M. (2004). How does preference reversal appear and disappear? Effects of the evaluation mode. *Journal of Behavioral Decision Making*, *17*, 395–408.

Soman, D., Ainslie, G., Frederick, S., Xiuping, L., Moreau, P., Mitchell, A., et al. (2005). (2005). The psychology of intertemporal discounting: Why are distant events valued more than proximal events? *Marketing Letters*, 347–360.

Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox. *Bookings Papers on Economic Activity*, 1–87.

Tehila, K., & Ritov, I. (2005). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, *97*, 106–116.

Van Boven, L., Loewenstein, G., & Dunning, D. (2005). The illusion of courage in social predictions: Underestimating the impact of fear of embarrassment on other people. *Organizational Behavior and Human Decision Processes*, *96*, 130–141.

Veenhoven, R. (1991). Is happiness relative? *Social Indicators Research*, *24*, 1–34.

Willemsen, M.C., & Keren, G. (2004). The role of negative features in joint and separate evaluation. *Journal of Behavioral Decision Making*, *17*, 313–329.

Wilson, T.D., & Gilbert, D. (2003). Affective forecasting. *Advances in Experimental Social Psychology*, *35*, 345–411.

Yeung, C., & Soman, D. (2005). Attribute evaluability and the range effect. *Journal of Consumer Research*, *32*, 363–369.

Yeung, C., & Soman, D. (2007). The duration heuristic. *Journal of Consumer Research*, *34*, 315–326.

Zhang, J., & Hsee, C.K. (2009). *Distinction bias: The moderating role of attribute evaluability*. Unpublished manuscript.

Zikmund-Fisher, B.J., Fagerlin, A., & Ubel, P.A. (2004). "Is 28% good or bad?" Evaluability and preference reversals in health care decisions. *Medical Decision Making*, *24*, 142–148.