Magnitude, Time, and Risk Differ Similarly between Joint and Single Evaluations

CHRISTOPHER K. HSEE
JIAO ZHANG
LIANGYAN WANG
SHIRLEY ZHANG

Arguably, all choice options involve three basic attributes: magnitude (outcome size), time (of occurrence), and probability (of occurrence) and are evaluated in one of two basic evaluation modes: JE (joint evaluation, involving comparison of multiple options) and SE (single evaluation, without comparison). This research explores how reactions to the three basic attributes (and their associated functions—utility, time discounting, and probability weighting) vary between the two basic evaluation modes. Nine studies, tapping diverse contexts, yield two general results: first, for all these attributes, people are more sensitive to variations near endpoints (zero magnitude, no delay, and 0% or 100% probability) than in other regions, and this differential sensitivity is more pronounced in SE than in JE. Second, when faced with options involving a trade-off between magnitude and time (delay) or between magnitude and probability (risk), people are both more delay averse and more risk averse in SE than in JE.

How much would people value receiving a $100 gift coupon for certain today? What about receiving a $150 gift coupon for certain in 2 weeks? What about an 80% chance to receive a $150 gift coupon today? Will people value these options differently if each option is presented and evaluated in isolation than if these options are juxtaposed and evaluated in comparison?

Christopher K. Hsee (chris.hsee@chicagobooth.edu) is Theodore O. Yntema Professor of Behavioral Science and Marketing, University of Chicago, Chicago, IL 60637. Jiao Zhang (jiaozhang@miami.edu) is assistant professor of marketing, University of Miami, Coral Gables, FL 33124. Liangyan Wang (wly@sjtu.edu.cn) is associate professor of marketing, and Shirley Zhang (shirleyzyw@hotmail.com) was a marketing student at Shanghai Jiaotong University (SJTU), Shanghai, China 200052. Correspondence: Christopher K. Hsee or Jiao Zhang. The authors acknowledge the following individuals (in alphabetical order of their last names) for helpful input at various stages of this project: Dan Bartels, David Levari, Tony Lian, Daniel Read, Xueer Yu, the editor, associate editor, and reviewers. The authors also thank their respective schools, the Templeton Foundation, and the National Natural Science Foundation of China for research support. Supplemental materials, describing additional statistical analyses comparing the curvatures of JE and SE functions, are available online.

Mary Frances Luce served as editor and Joel Huber served as associate editor for this article.

Electronically published January 23, 2013

Arguably all choice options involve three basic attributes —magnitude (size of outcome), time (time of occurrence), and probability (likelihood of occurrence). For example, the option "receiving a $150 gift coupon for certain in 2 weeks" has a magnitude of $150, a probability of 100%, and a time (delay) of 2 weeks. The option "an 80% chance to receive a $150 gift coupon today" has a magnitude of $150, a probability of 80%, and a time (delay) of zero. Magnitude does not have to be monetary. For example, the option "a 20% chance to save 40 dolphins in 3 months" has a magnitude of 40 dolphins.

Arguably all evaluations of choice options occur in one of two modes—joint evaluation (JE) and single evaluation (SE)—or in some combination of the two. By JE, we mean that two or more options are juxtaposed and evaluated comparatively; by SE, we mean that only one option is presented and evaluated in isolation (Hsee 1996; Hsee and Zhang 2010). Evaluation mode is a fundamental characteristic of judgment and choice and can be likened to the rain-shine dimension of weather. Just as any day can be characterized as rainy or sunny, any judgment or choice can be characterized as made in JE or SE. Sometimes we encounter one option at a time and have to decide whether to accept it or how much effort to expend to accept it; this is an example of SE. Sometimes we encounter multiple possible options at the same time and have to decide which one to accept, or how much effort to expend to acquire one option versus

another option; this is an example of JE. Evaluation mode may fall between JE and SE, such as when there is one highly salient option along with several less salient options. In this article, we look at situations that are clearly JE or SE.

Reactions to magnitude, time, and probability have been extensively studied in the existing literature (e.g., Frederick, Loewenstein, and O'Donoghue 2002; Gonzalez and Wu 1999; Kahneman and Tversky 1979). Evaluation mode has already been well studied (e.g., Bazerman et al. 1999; Bleichrodt and Prades 2009; Burson, Larrick, and Lynch 2009; Chatterjee, Heath, and Min 2009; Hsee 1996; Hsee and Leclerc 1998; Hsee et al. 1999; Kogut and Ritov 2005; Laran and Wilcox 2011; List 2002; Nowlis and Simonson 1997; Okada 2005; Paharia et al. 2009; Schmeltzer, Caverni, and Warglien 2004). Nevertheless, the relationship between the two topics has not been well studied.

In the current research, we seek to make two contributions. First, we study how reactions to the three basic attributes vary between the two basic evaluation modes. Traditionally, reactions to the three attributes have largely been studied separately, each having its own functions and its own factors. Reactions to magnitude are usually modeled by utility and value functions (e.g., Kahneman and Tversky 1979); factors identified as influencing magnitude valuation include perceived meaning of zero (Gneezy and Rustichini 2000; Palmeira 2011; Shampanier, Mazer, and Ariely 2007) and affect richness of the stimulus (Hsee and Rottenstreich 2004). Reactions to time are typically modeled by time-discounting functions; factors identified as influencing time discounting include presentation format (Ebert and Prelec 2007; Malkoc and Zauberman 2006; Read 2001; Read et al. 2005; Zauberman et al. 2009), perceived connectedness between the current and future self (Bartels and Rips 2010; Bartels and Urminsky 2011; Hershfield et al. 2011), subjective time perception (Zauberman and Lynch 2005; Zauberman et al. 2009), visceral states (Loewenstein 1996; Van den Bergh et al. 2008), and age and intelligence (Read and Read 2004; Shamosh et al. 2008). Reactions to probability are typically modeled by probability weighting functions; factors identified as influencing probability weighting include source of uncertainty (Kilka and Weber 2001), size of outcome (Prelec and Loewenstein 1991; Weber and Chapman 2005), regulatory focus (Kluger et al. 2004), and individual differences (Chen and Jia 2005). Prior research has also examined the connection between time discounting and probability weighting and tried to identify factors that influence both (e.g., Baucells and Heukamp 2012; Bommier 2006; Prelec and Loewenstein 1991; Rachlin, Brown, and Cross 2000). Joining this effort, we propose and demonstrate that evaluation mode is a common factor that influences reactions to all of the three basic sets of functions; for each of the three attributes, people are more sensitive to variations near endpoints than elsewhere, and this differential sensitivity is more pronounced in SE than in JE.

The second contribution we seek to make is an extension of the first; we study how preferences for options involving two of the three basic attributes vary between JE and SE. These two-attribute trade-off preferences include time preference, involving a trade-off between magnitude and time, and risk preference, involving a trade-off between magnitude and probability. Time preference and risk preference are among the most common trade-offs people encounter in real life. Examples of time preference include whether to take one bird today or wait for two birds tomorrow, whether to cash a check now or wait a year for a higher amount, and whether to buy an 8-gigabyte Kindle now or wait 3 months until a 16-gigabyte edition is released. Examples of risk preference include whether to accept one bird in the hand or seek two birds in the bush, whether to invest money in a savings account with a lower fixed rate or in a mutual fund with a higher but risky rate, and whether to buy a new bicycle with known quality or a used bicycle with risky but potentially higher quality. Time preference and risk preference have received substantial attention in the existing literature (on time preference: e.g., Frederick et al. 2002; Kable and Glimcher 2010; Mischel, Shoda, and Rodriguez 1989; Read et al. 2005; on risk preference: e.g., Gneezy, List, and Wu 2006; Gonzalez and Wu 1999; Prelec 1998; Slovic 1987). We seek to contribute to the existing literature by studying how these preferences vary between JE and SE. In particular, we propose and demonstrate that, in time preference, people are more delay averse in SE than in JE and that, in risk preference, people are more risk averse in SE than in JE.

REACTIONS TO MAGNITUDE, TIME, AND PROBABILITY

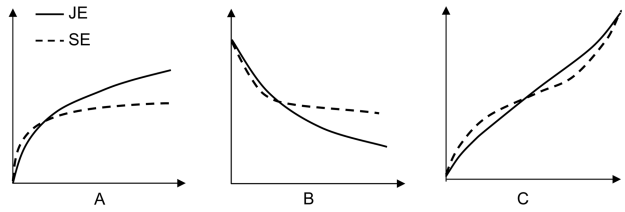
In this section, we examine how people's reactions to a single attribute—magnitude, time, or probability—vary between JE and SE. Namely, we examine how each of the three basic functions—magnitude valuation, time discounting, and probability weighting functions—varies between JE and SE.

Existing research suggests that the three basic functions share a common pattern: they are all steeper near the endpoints and flatter away from the endpoints. Generally speaking, the endpoint in magnitude valuation is zero magnitude (e.g., $0); in temporal discounting, zero delay (immediate); and in probability weighting, certainty (100%) or impossibility (0%). The nonlinearity of these functions is multiply determined, including diminishing marginal utility for magnitudes, uncertainty about the future in time discounting, and general psychophysical diminishing marginal sensitivity in all the three cases.

In the present research, we do not intend to explain why these functions are generally flatter when away from the endpoints than when near the endpoints. Instead, we posit that when away from the endpoints, but not when near the endpoints, all of these functions are even flatter in SE than in JE (see fig. 1). This idea has been proposed by Hsee and Zhang (2010) but has not been tested. The current research elaborates on this proposition and provides empirical support. Furthermore, the current research also examines the

FIGURE 1

HYPOTHESIZED MAGNITUDE VALUATION, TIME DISCOUNTING, AND PROBABILITY WEIGHTING FUNCTIONS UNDER JE AND SE



NOTE.—Adapted from figure 3 in Hsee and Zhang 2010.

implications of the proposition for trade-off preferences, including risk preference and time preference.

Why do the JE and SE functions differ? In SE, people are presented with only one value, and they tend to evaluate it by its category rather than by its magnitude because category is generally easier to evaluate than magnitude (Hsee and Zhang 2004, 2010). In JE, people are presented with multiple values and can directly compare and evaluate them by their (relative) magnitude, rather than by just their category. Thus, people in SE are sensitive to the category in which a given value falls and relatively insensitive to its magnitude, whereas people in JE are more sensitive to its magnitude. Suppose that x_1 and x_2 belong to two different categories on a given attribute, and x_2 and x_3 belong to the same category on the same attribute, and objectively $x_3 - x_2 = x_2 - x_1 > 0$. Then x_1 and x_2 will be evaluated differently in both SE and JE, whereas x_2 and x_3 will be evaluated similarly in SE but differently in JE.

Categories depend on salient reference points; in other words, salient reference points on the target attribute divide its values into different categories. In magnitude valuation, the most salient reference point is zero; for example, whether the price of a product involves no (zero) discount or some discount off its manufacturer's suggested retail price (MSRP). Therefore, if one option involves a magnitude of zero and another a magnitude of nonzero (e.g., if one store offers no discount on a certain product and another offers some discount on the same product), the two magnitudes are categorically different and will be evaluated differently in both SE and JE. However, if both magnitudes are nonzero and one is greater than the other (e.g., if one store offers a $10 discount and another offers a $20 discount), they are only incrementally different and will be evaluated similarly in SE but still differently in JE. This explains why magnitude valuation functions are similar between SE and JE around zero but different away from zero (fig. 1*A*). Specifically, we propose the following hypothesis:

H1: (concerning magnitude valuation): Around the zero-magnitude point, magnitude valuation functions are as steep (or steeper) in SE as (than) in JE, and away from the zero-magnitude point, magnitude valuation functions are flatter in SE than in JE.

We propose that the analysis for magnitude valuation functions applies also to time-discounting and probability weighting functions. In time discounting, the most salient reference point is zero delay (immediate; Scholten and Read 2006). Therefore, if one option involves zero delay and another option involves some delay, the two options are categorically different and will be evaluated differently in both SE and JE. If one option involves a shorter delay and another option involves a longer delay, then the two options are only incrementally different and will be evaluated similarly in SE yet still differently in JE. This leads to our second hypothesis (fig. 1*B*):

H2: (concerning time discounting): Around the zero-delay point, time-discounting functions are as steep (or steeper) in SE as (than) in JE, and away from the zero-delay point, time-discounting functions are flatter in SE than in JE.

Likewise, in probability weighting, the most salient reference points are certainty (100%) and impossibility (0%). Thus, if one option is certain (100%) and another option is probabilistic, or one option is impossible (0%) and another is probabilistic, then the two options are categorically different and will be evaluated differently in both SE and JE. If both options are probabilistic, with one being more probable than the other, the two options are only incrementally different, and they will be evaluated similarly in SE (Kunreuther, Novemsky, and Kahneman 2001) yet still differently in JE. This leads to our third hypothesis (fig. 1*C*):

H3: (concerning probability weighting): Around certainty (100%) and impossibility (0%), probability weighting functions are as steep (or steeper) in SE as (than) in JE, and away from certainty and impossibility, probability weighting functions are flatter in SE than in JE.

Of these three hypotheses, the hypothesis about magnitude valuation (hypothesis 1) has already been suggested in earlier writings (Hsee and Zhang 2004; Hsee et al. 1999). The present research extends the hypothesis about magnitude valuation to time discounting (hypothesis 2) and probability weighting (hypothesis 3) and shows that these three functions behave similarly with respect to JE versus SE.

TRADE-OFF PREFERENCES

In the previous section, we focused on how people's reactions to a single attribute (magnitude, time, or probability) vary between JE and SE. In this section, we examine how preferences for options involving two attributes vary between JE and SE. For example, between $100 now and $150 in a month, which one is more appealing in JE and which is more appealing in SE? Between $100 for sure and an

80% chance at $150, which is more appealing in JE and which is more appealing in SE? The first question concerns how time preference varies between JE and SE; the second question concerns how risk preference varies between JE and SE.

According to our theory (hypotheses 1–3), as long as two options involve a trade-off between any two of the three basic attributes (magnitude, time, and probability), and one of the two attributes entails a categorical difference and the other attribute entails an incremental difference, then preferences for these options may reverse between JE and SE. That is because people are more sensitive to the incremental difference in JE than in SE. To illustrate, consider two options: A = $100 now versus B = $150 in a month. Note that the difference in magnitude ($100 vs. $150) is incremental, and the difference in time (now vs. delayed) is categorical. According to hypothesis 1, people in JE would be more sensitive to the difference between $100 and $150 than people in SE, and, according to hypothesis 2, people in SE would be equally sensitive (if not more sensitive) to the difference between no delay and a delay than people in JE. Therefore, we expect that, relative to B (the larger delayed outcome), A (the smaller immediate outcome) will be favored more in SE than in JE. The same can be said of trade-offs between incremental magnitude (smaller vs. larger) and categorical probability (certain vs. uncertain), for example, between $100 for sure and an 80% chance at $150. These analyses lead to the following hypotheses, one about time preferences and one about risk preferences:

H4: (concerning trade-offs between incremental magnitude and categorical time): If one option entails a larger delayed gain and another option entails a smaller immediate (no-delay) gain, then, relative to the first option, the second option will be favored more in SE than in JE.

H5: (concerning trade-offs between incremental magnitude and categorical probability): If one option entails a larger risky gain and another option entails a smaller riskless gain, then, relative to the first option, the second option will be favored more in SE than in JE.

Hypothesis 4 suggests that, for time preferences, people are more delay averse in SE than in JE. Hypothesis 5 suggests that, for risk preferences, people are more risk averse in SE than in JE. (Excessive aversion to delays has been termed the present bias [Laibson 1997]; using this term, hypothesis 4 suggests that the present bias is more prominent in SE than in JE.)

In theory, there are eight possible categorical-versus-incremental difference trade-offs involving two of the three basic attributes (magnitude, time, and probability). In reality, some of these trade-offs are not meaningful. For example, it is not meaningful for an option to have a nonzero magnitude and a zero probability, or a nonzero probability and a zero magnitude. There are only four meaningful trade-offs—between incremental magnitude and categorical time, between incremental magnitude and categorical probability, between incremental time and categorical probability, and between categorical time and incremental probability. Of these four, the first two have already been discussed in hypotheses 4 and 5. Applying the same logic behind hypotheses 4 and 5, we propose the following two hypotheses regarding the remaining two trade-offs, both about time and probability:

H6: (concerning trade-offs between incremental time and categorical probability): If one option entails a less delayed risky gain and another option entails a more delayed riskless gain, then, relative to the first option, the second will be favored more in SE than in JE.

H7: (concerning trade-offs between categorical time and incremental probability): If one option entails a delayed but less risky gain and another option entails an immediate (no-delay) but riskier gain, then, relative to the first option, the second option will be favored more in SE than in JE.

The existing literature has documented various forms of preference reversals, including preference reversals related to risk preference. For example, in the classic choice-matching and choice-pricing paradigms (Lichtenstein and Slovic 1971; Tversky, Sattath, and Slovic 1988), choosers tend to weigh probability more heavily whereas matchers and price setters tend to weigh outcome magnitude more heavily. Although these paradigms and the current research are all about preference reversals, they have different foci. The classic paradigms are concerned with reversals between different response scales—for example, choice versus matching—and the alternatives are typically presented together, as in JE. In contrast, the present research is about reversals between different evaluation modes, and the response scale is held constant across conditions, for example, always willingness to pay. The present research is a first attempt to systematically demonstrate JE-SE reversals in time preference, risk preference, and trade-off preferences between time and risk.

We now report a series of studies that tested our hypotheses. Of these studies, the first four tested the three initial hypotheses regarding single attributes (hypotheses 1–3), and the remaining studies tested the subsequent hypotheses regarding trade-off preferences involving two attributes (hypotheses 4–7). These studies tapped a wide range of domains, and two of the studies (6 and 8) involved real consequences for participants.

STUDY 1

Methods

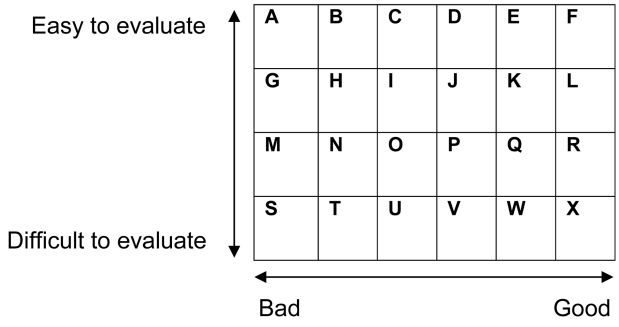
The main purpose of this study was to test hypothesis 1, regarding magnitude valuation. A secondary purpose was to explore the underlying mechanism. This study examined participants' willingness to donate to a polar bear rescue program that had three alternative achievement records—-

having saved 0, 100, or 200 polar bears in the past 5 years. To explore the underlying mechanism, the study included an evaluation grid, as illustrated in figure 2, and asked the participants to select a letter from the grid before indicating their willingness to donate. The choice of any letter in the evaluation grid revealed a participant's evaluations on two dimensions at the same time: desirability (bad vs. good) and evaluability (difficult vs. easy to evaluate the desirability). We coded the evaluations on the desirability dimension on a 1 (bad) to 6 (good) scale, and the evaluations on the evaluability dimension on a 1 (difficult to evaluate) to 4 (easy to evaluate) scale. For example, the letter K reflects a rating of 5 on the desirability dimension and a rating of 3 on the evaluability dimension.

Participants were 538 individuals (55% female) recruited through an online survey service in the United States. In a questionnaire, the participants were first presented and made familiar with the evaluation grid. Then they were asked to imagine that they were in a zoo when a fund-raiser from a nonprofit organization approached them to solicit cash donations to support an ongoing program to save polar bears, and that they had recently read a report from a trustworthy third party about the history of the program. Participants were randomly assigned to one of six conditions: three SE conditions and three JE conditions. In the three SE conditions, participants were told that, according to the report, the program had received generous support from the public in the past 5 years and had saved 0, 100, or 200 polar bears, respectively, during the period. They were asked to evaluate the ability of the program to save polar bears by choosing a letter from the evaluation grid. They were then asked to assume that they had $100 cash on them at the moment and to indicate how much they would donate to support the program by writing down an amount between $0 and $100. In the three JE conditions, participants received similar information and instructions as those in the SE conditions except that, in each JE condition, they read two of the three alternative scenarios and were asked to evaluate the program using the evaluation grid and to indicate their willingness to donate in each of the two scenarios. The order of the two scenarios was counterbalanced.

FIGURE 2

THE EVALUATION GRID USED IN STUDY 1



Results and Discussion

Willingness to Donate. The results are summarized in table 1. Hypothesis 1 predicts that respondents would be similarly sensitive in JE and in SE to whether the program had saved 0 or 100 polar bears and would be more sensitive in JE than in SE to whether the program had saved 100 or 200 polar bears. To test these predictions, we performed two hybrid *t*-tests as proposed by Hsee (1996), one for 0 versus 100 bears and one for 100 versus 200 bears. According to the first hybrid *t*-test, respondents were indeed similarly sensitive in JE and SE to whether the program had saved 0 or 100 bears ($t(300) = 1.62$, $p > .10$, NS); in both evaluation modes, respondents were willing to donate more in the 100-bear condition than in the 0-bear condition (in SE, $M_{100\ \text{bears}} = 18.05$ vs. $M_{0\ \text{bear}} = .81$, $t(218) = 6.28$, $p < .0001$; in JE, $M_{100\ \text{bears}} = 24.02$ vs. $M_{0\ \text{bear}} = .66$, $t(82) = 7.95$, $p < .0001$). According to the second hybrid *t*-test, respondents were indeed more sensitive in JE than in SE to whether the program had saved 100 or 200 bears ($t(277) = 4.91$, $p < .0001$); in SE, respondents were unwilling to donate more (in fact, they were willing to donate less) in the 200-bear condition than in the 100-bear condition ($M_{200\ \text{bears}} = 11.80$ vs. $M_{100\ \text{bears}} = 18.05$, $t(204) = -1.93$, $p = .056$), whereas in JE, respondents were willing to donate more in the 200-bear condition than in the 100-bear condition ($M_{200\ \text{bears}} = 25.19$ vs. $M_{100\ \text{bears}} = 13.73$, $t(73) = 5.25$, $p < .0001$). We also performed a hybrid *t*-test for 0 versus 200 bears and found that respondents were more sensitive in JE than in SE to whether the program had saved 0 or 200 bears ($t(252) = 3.96$, $p < .0001$); in both evaluation modes, respondents were willing to donate more in the 200-bear condition than in the 0-bear condition (in SE, $M_{200\ \text{bears}} = 11.80$ vs. $M_{0\ \text{bear}} = .81$, $t(170) = 6.95$, $p < .0001$; in JE, $M_{200\ \text{bears}} = 26.70$ vs. $M_{0\ \text{bear}} = 1.17$, $t(82) = 7.85$, $p < .0001$). Together, these results supported our prediction that magnitude valuation functions vary depending on evaluation modes—similarly steep in JE and in SE when near the endpoint, and steeper in JE than in SE when away from the endpoint.

Desirability and Evaluability. These were ratings on the evaluation grid; they provided additional insight into how participants responded to the magnitude variable (number of polar bears saved). We first report the relationship between the desirability ratings and the magnitude valuable and then report the relationship between the evaluability ratings and the magnitude variable.

Regarding the relationship between the desirability ratings and the magnitude variable, we predicted that it would closely match the relationship between willingness to donate and the magnitude variable. This was indeed the case. Over the 0-to-100-bear range, desirability ratings increased by similar amounts in JE and in SE ($t(300) = .69$, NS, using the hybrid *t*-test); in both evaluation modes, respondents found the program more desirable if it had saved 100 polar bears than if it had saved 0 (in SE, $M_{100\ \text{bears}} = 4.45$ vs.

TABLE 1

RESULTS OF ALL OF THE STUDIES INVOLVING A SINGLE ATTRIBUTE

Study	JE mean (SD)	SE mean (SD)
Study 1 (willingness to donate to a polar bear rescue program, in $):		
0 and 100 polar bears	.66 (5.51)[a]	.81 (3.12)[a]
	24.02 (26.79)[b]	18.05 (26.33)[b]
0 and 200 polar bears	1.17 (5.48)[a]	.81 (3.12)[a]
	26.70 (30.03)[b]	11.80 (14.87)[b]
100 and 200 polar bears	13.73 (18.03)[a]	18.05 (26.33)[a]
	25.19 (27.93)[b]	11.80 (14.87)[b]
Study 2 (willingness to tutor for an amount to be paid at various time points, in ¥):		
Now	140.53 (64.34)[a]	120.73 (48.13)[a]
1 month	109.74 (64.79)[b]	78.75 (64.34)[b]
3 months	58.42 (64.66)[c]	80.49 (68.55)[b]
Study 3a (willingness to wait for a chance to acquire a certificate, in minutes):		
100%	85.52 (42.94)[a]	66.08 (40.68)[a]
90%	69.72 (39.80)[b]	53.81 (35.44)[b]
80%	57.02 (41.51)[c]	60.30 (37.02)[a,b]
Study 3b (willingness to wait for a chance to acquire a certificate, in minutes):		
0%	.59 (2.30)[a]	3.99 (9.74)[a]
10%	17.31 (25.45)[b]	26.69 (30.84)[b]
20%	31.63 (34.28)[c]	26.29 (25.17)[b]
Additional study reported in General Discussion (willingness to pay for cactus of various diameters, in $):		
Reference = 60 centimeters:		
60 centimeters	225.91 (147.02)[a]	301.25 (181.38)[a]
70 centimeters	322.27 (157.39)[b]	420.96 (222.34)[b]
80 centimeters	456.36 (294.80)[c]	407.04 (154.57)[a,b]
Reference = 80 centimeters:		
60 centimeters	180.60 (94.65)[a]	228.63 (111.89)[a]
70 centimeters	309.60 (117.42)[b]	243.70 (132.06)[a,b]
80 centimeters	467.20 (112.45)[c]	318.89 (163.81)[b]

NOTE.—The superscripts indicate whether or not means in the same evaluation mode (JE or SE) condition of each study are significantly different from each other. Means with different superscripts are significantly different from each other; means with the same superscript are not.

$M_{0\ \text{bear}} = 1.30$, $t(218) = 19.83$, $p < .0001$; in JE, $M_{100\ \text{bears}} = 4.42$ vs. $M_{0\ \text{bear}} = 1.12$, $t(82) = 20.44$, $p < .0001$). Over the 100-to-200-bear range, desirability ratings increased more in JE than in SE ($t(277) = 6.56$, $p < .0001$ using the hybrid t-test); in SE, respondents found the program equally desirable whether it had saved 200 or 100 polar bears ($M_{200\ \text{bears}} = 4.16$ vs. $M_{100\ \text{bears}} = 4.45$, $t(204) = -1.42$, NS), whereas in JE, respondents found the program more desirable if it had saved 200 polar bears than if it had saved 100 polar bears ($M_{200\ \text{bears}} = 4.85$ vs. $M_{100\ \text{bears}} = 3.59$, $t(73) = 11.54$, $p < .0001$). Over the 0-to-200-bear range, desirability ratings also increased more in JE than in SE ($t(252) = 3.45$, $p < .001$ using the hybrid t-test); in both evaluation modes, respondents found the program more desirable if it had saved 200 polar bears than if it had saved 0 (in SE, $M_{200\ \text{bears}} = 4.16$ vs. $M_{0\ \text{bear}} = 1.30$, $t(170) = 15.12$, $p < .0001$; in JE, $M_{200\ \text{bears}} = 4.94$ vs. $M_{0\ \text{bear}} = 1.18$, $t(82) = 22.31$, $p < .0001$).

More pertinent to our theory was the relationship between the evaluability ratings and the magnitude variable. Based on our theory concerning the role of the endpoint (zero magnitude), we made two predictions. First, evaluability for zero magnitude (0 bear) would be higher than evaluability for nonzero magnitudes (100 or 200 bears). Second, the difference in evaluability between the zero and nonzero magnitudes would be larger in SE than in JE. Consistent with our first prediction, in both evaluation modes, evaluability ratings were indeed higher for the zero magnitude than for the nonzero magnitude conditions (in SE, $M_{\text{zero mag.}} = 3.56$ vs. $M_{\text{nonzero mag.}} = 2.90$, $t(297) = 5.75$, $p < .0001$, by combining the 100-bear-SE and the 200-bear-SE conditions and comparing the combined data with the 0-bear-SE condition; in JE, $M_{\text{zero mag.}} = 3.48$ vs. $M_{\text{nonzero mag.}} = 3.11$, $t(165) = 3.19$, $p < .01$, by comparing the zero and the nonzero magnitude results in the 0-to-100-bear-JE and the 0-to-200-bear-JE conditions). Consistent with our second prediction, the difference in evaluability ratings between the zero magnitude and the nonzero magnitude conditions was (marginally significantly) greater in SE than in JE ($t(462) = 1.89$, $p = .059$, using the hybrid t-test). These results suggested that it was easier to evaluate the endpoint (zero magnitude) than to evaluate the other values and that joint eval-

uation reduced the difference in evaluability between the endpoint and the other values.

STUDY 2

Methods

Designed to test hypothesis 2 regarding the time-discounting function, this study examined students' willingness to work for a fixed fee that would be paid at one of three possible time points: immediately, in 1 month, and in 3 months. To test for the generality of our findings, we used a somewhat different format for JE in study 2 (and study 3) than for study 1. In the JE condition of study 1, the three magnitude levels were presented in pairs—two at a time. In the JE condition of study 2 (and study 3), the three levels were presented together, with order randomized.

Participants were 160 students (36% female) from Shanghai Jiaotong University in China. In a questionnaire, they were asked to indicate the maximum amount of time they were willing to spend tutoring a middle-school student to help her pass a final exam and earn ¥100. Participants were assigned to one of four conditions: one JE and three SE conditions. Those in the JE condition were asked to consider three alternative scenarios of when they would be paid (presented side by side, order randomized): immediately, in 1 month, or in 3 months. They were asked to indicate their willingness to work in each of these three scenarios by writing down a number between 0 and 240 minutes. Participants in each SE condition received the same instructions except that they saw only one of the three scenarios and indicated their willingness to work in that scenario.

Results and Discussion

The results, summarized in table 1, were consistent with hypothesis 2 and the pattern in figure 1*B*. We performed two hybrid *t*-tests, one for no delay versus 1-month delay and one for 1-month delay versus 3-month delay. According to the first hybrid *t*-test, respondents were similarly sensitive in JE and in SE to whether there was a 1-month delay or no delay ($t(116) = .68$, NS); in both evaluation modes, respondents were willing to work more in the no-delay condition than in the 1-month-delay condition (in SE, $M_{\text{no delay}} = 120.73$ vs. $M_{1\text{ mo. delay}} = 78.75$, $t(79) = 3.33$, $p < .01$; in JE, $M_{\text{no delay}} = 140.53$ vs. $M_{1\text{ mo. delay}} = 109.74$, $t(37) = 2.94$, $p < .01$). According to the second hybrid *t*-test, respondents were more sensitive in JE than in SE to whether there was a 1-month delay or a 3-month delay ($t(116) = 2.94$, $p < .01$); in SE, respondents were willing to work about the same amount in the 1-month-delay condition as in the 3-month-delay condition ($M_{1\text{ mo. delay}} = 78.75$ vs. $M_{3\text{ mo. delay}} = 80.49$, $t(79) = -.12$, NS), whereas in JE, respondents were willing to work more in the 1-month-delay condition than in the 3-month-delay condition ($M_{1\text{ mo. delay}} = 109.74$ vs. $M_{3\text{ mo. delay}} = 58.42$, $t(37) = 4.95$, $p < .0001$). These results supported our hypothesis that the shape of the time-discounting function depends on evaluation modes—similarly steep in JE and SE near the zero-delay point, and flatter in SE than in JE away from the zero-delay point.

STUDY 3A

Methods

Designed to test hypothesis 3 regarding probability weighting around the certainty end (100%), this study examined students' willingness to wait for a chance to acquire a gift certificate with three chance (probability) levels: 100%, 90%, and 80%.

Participants were 333 students, of whom 172 (48% female) were from the University of Miami and 161 (47% females) from Shanghai Jiaotong University in China. They were asked to imagine that a new department store had just opened in their neighborhood and that, as a promotion, the store was giving away gift certificates. Each certificate was worth $50 (or ¥100 in the version for participants in China) and was redeemable at the store immediately. To have a chance to acquire a gift certificate, one had to stand in line. Participants were assigned to either a JE condition or one of three SE conditions. Those in the JE condition were asked to consider three alternative scenarios of their likelihood to get the gift certificate (presented side by side, order randomized): 100%, 90%, and 80%. They were asked to assume that they had up to 120 minutes of free time and to indicate the maximum amount of time (in minutes) they were willing to wait. Participants in each SE condition received the same instructions, except that they saw only one of the three scenarios and indicated their willingness to wait in that scenario.

Results and Discussion

The results, summarized in table 1, supported hypothesis 3 and were consistent with the pattern in figure 1*C*. Again, we performed two hybrid *t*-tests, one for 100% versus 90% and one for 90% versus 80%. According to the first hybrid *t*-test, respondents were similarly sensitive in JE and in SE to whether the chance was 100% or 90% ($t(248) = -.34$, NS); in both evaluation modes, respondents were willing to wait longer in the 100% probability condition than in the 90% probability condition (in SE, $M_{100\%} = 66.08$ vs. $M_{90\%} = 53.81$, $t(159) = 2.03$, $p < .05$; in JE, $M_{100\%} = 85.52$ vs. $M_{90\%} = 69.72$, $t(89) = 6.95$, $p < .0001$). According to the second hybrid *t*-test, respondents were more sensitive in JE than in SE to whether the chance was 90% or 80% ($t(246) = 2.50$, $p = .01$); in SE, respondents were willing to wait as long in the 90% probability condition as in the 80% probability condition ($M_{90\%} = 53.81$ vs. $M_{80\%} = 60.30$, $t(157) = -1.13$, NS), whereas in JE, respondents were willing to wait longer in the 90% probability condition than in the 80% probability condition ($M_{90\%} = 69.72$ vs. $M_{80\%} = 57.02$, $t(89) = 9.51$, $p < .0001$). In support of hypothesis 3, the results suggested that respondents were similarly sensitive in SE and JE to whether acquiring a certificate was guaranteed or not, yet when acquisition was not guaranteed,

they were less sensitive in SE than in JE to how big the chance was.

STUDY 3B

Methods

Study 3b was identical to study 3a except that it examined the impossibility (0%) end of probability with three chance (probability) levels: 0%, 10%, and 20%. Participants were 336 students, of whom 182 (50% female) were from the University of Miami and 154 (48% females) from Shanghai Jiaotong University in China.

Results and Discussion

As in study 3a, we performed two hybrid *t*-tests, one for 0% versus 10% and one for 10% versus 20%. According to the first hybrid *t*-test, respondents were similarly sensitive in JE and in SE to whether the probability was 0% or 10% ($t(250) = 1.31$, NS); in both evaluation modes, respondents were willing to wait longer in the 10% probability condition than in the 0% probability condition (in SE, $M_{10\%} = 26.69$ vs. $M_{0\%} = 3.99$, $t(164) = 6.55$, $p < .0001$; in JE, $M_{10\%} = 17.31$ vs. $M_{0\%} = .59$, $t(86) = 6.14$, $p < .0001$). According to the second hybrid *t*-test, respondents were more sensitive in JE than in SE to whether the probability was 10% or 20% ($t(245) = 3.07$, $p < .01$); in SE, respondents were willing to wait as long in the 20% probability condition as in the 10% probability condition ($M_{20\%} = 26.29$ vs. $M_{10\%} = 26.69$, $t(159) = -.09$, NS), whereas in JE, respondents were willing to wait longer in the 20% probability condition than in the 10% probability condition ($M_{20\%} = 31.63$ vs. $M_{10\%} = 17.31$, $t(86) = 8.14$, $p < .0001$). Consistent with hypothesis 3, these results suggested that the respondents were similarly sensitive in SE and JE to whether there was a chance to acquire a certificate, yet when there was a chance, they were less sensitive in SE than in JE to how big the chance was.

Put together, the results of the studies reported so far supported our proposition that, in JE, people are similarly sensitive to categorical and incremental differences whereas, in SE, people are more sensitive to categorical differences than to incremental differences. Studies 1–3b tested hypotheses 1–3, each concerning a single attribute; the remaining studies tested hypotheses 4–7, each concerning a trade-off between two attributes.

STUDY 4

Methods

Study 4 tested hypothesis 4 concerning time preferences involving a trade-off between immediacy and magnitude. It examined respondents' willingness to pay for either a slower but immediately available Internet service (smaller immediate gain) or a faster but not immediately available Internet service (larger delayed gain).

Participants were 96 students (54% female) from Shanghai Jiaotong University in China. In a questionnaire, they were asked to imagine that they planned to subscribe to a 1-year Internet service plan. There were two options: one featured a 2 megabyte per second speed and would be installed immediately, and the other featured a 4 megabyte per second speed and would not be available for installation until 3 months later. In either case, the subscription fee had to be paid up front. Participants were randomly assigned to one of three conditions: one JE and two SE. Those in the JE condition saw both service options (presented side by side, order counterbalanced) and indicated their willingness to pay for each option. Participants in each SE condition saw only one of the two service options and indicated their willingness to pay for that option.

Results and Discussion

According to hypothesis 4, relative to larger delayed gains, smaller immediate gains will be favored more in SE than in JE. The results, summarized in table 2, indeed exhibited this pattern. A hybrid *t*-test revealed a significant JE-SE preference reversal ($t(93) = 2.56$, $p < .05$). In JE, respondents were willing to pay more for the delayed 4-megabyte service ($M_{\text{del. 4MB}} = 286.67$ vs. $M_{\text{immed. 2MB}} = 233.33$, $t(29) = 2.24$, $p < .05$) whereas, in SE, respondents were willing to pay somewhat (not significantly) more for the immediate 2-megabyte service ($M_{\text{del. 4MB}} = 272.73$ vs. $M_{\text{immed. 2MB}} = 336.36$, $t(64) = -1.63$, NS).

STUDY 5

Methods

Study 5 sought to replicate the finding of study 4 in a real rather than hypothetical decision context. It examined students' willingness to participate in two upcoming studies, one offering a smaller immediate payment and the other a larger delayed payment.

Participants were 68 students (56% female) from Shanghai Jiaotong University in China. They were randomly assigned to one of three conditions: one JE and two SE. Those in the JE condition were approached at public places on campus by a research assistant who was recruiting students to participate in two upcoming 10-minute paper-and-pencil studies on stress management. One of the upcoming studies would pay each participant ¥10 immediately afterward; the other upcoming study would pay each participant ¥25 but the payment would not be delivered until 2 months later. The participants were asked to indicate their willingness to participate in each of the upcoming studies by drawing a vertical line on a 5-inch scale anchored by "not willing at all" (left) and "very willing" (right). Responses were later converted into numbers; for example, if a participant drew a vertical line 2.5 inches from the left end of the scale, her response would be coded as 2.5. Participants in each SE condition received the same information except that they were told about only one of the upcoming studies and asked

TABLE 2

RESULTS OF ALL STUDIES INVOLVING TRADE-OFFS ALONG TWO ATTRIBUTES

Study	Options	JE mean (SD)	SE mean (SD)
Study 4 (willingness to pay for an Internet service, in ¥)	Larger, delayed	286.67 (159.16)[a]	272.73 (137.55)[a]
	Smaller, immediate	233.33 (88.41)[b]	336.36 (176.45)[a]
Study 5 (willingness to participate in a study, on 0–5 scale)	Larger, delayed	3.31 (1.06)[a]	2.24 (1.44)[a]
	Smaller, immediate	3.09 (.99)[a]	3.34 (1.30)[b]
Study 6 (willingness to pay for a rescue program, in million ¥)	Larger, risky	70.67 (26.51)[a]	70.00 (28.02)[a]
	Smaller, riskless	58.33 (30.07)[b]	79.13 (24.66)[a]
Study 7 (willingness to participate in a study, on 0–5 scale)	Larger, risky	2.69 (1.38)[a]	2.29 (1.30)[a]
	Smaller, riskless	2.47 (1.26)[a]	2.82 (1.07)[b]
Study 8 (willingness to pay for a repair service, in ¥)	Less delayed; risky	126.47 (81.24)[a]	110.00 (77.94)[a]
	More delayed; riskless	91.18 (90.55)[a]	187.50 (134.78)[b]
Study 9 (willingness to participate in a study, on 1–7 scale)	More likely; delayed	5.97 (1.17)[a]	4.10 (2.07)[a]
	Less likely; immediate	5.35 (1.25)[b]	5.13 (1.61)[b]

NOTE.—The superscripts indicate whether or not means in the same evaluation mode (JE or SE) condition of each study are significantly different from each other. Means with different superscripts are significantly different from each other; means with the same superscript are not.

to indicate their willingness to participate in that study. All participants were told that those who expressed the greatest willingness would be contacted first to participate and that, if contacted, they would be obligated to participate. About 10 participants in each condition who expressed the strongest willingness to participate in that study were indeed contacted later and paid accordingly after their participation.

Results and Discussion

The results, summarized in table 2, replicated those of study 4. A hybrid t-test revealed a significant JE-SE preference reversal ($t(65) = 2.80$, $p < .01$). In JE, respondents were somewhat (not significantly) more willing to participate in the upcoming study that offered a larger delayed payment ($M_{\text{del. ¥25}} = 3.31$ vs. $M_{\text{immed. ¥10}} = 3.09$, $t(25) = 1.02$, NS) whereas, in SE, respondents were more willing to participate in the upcoming study that offered a smaller immediate payment ($M_{\text{del. ¥25}} = 2.24$ vs. $M_{\text{immed. ¥10}} = 3.34$, $t(40) = -2.62$, $p < .05$). Together, studies 4 and 5 supported hypothesis 4, showing that, relative to an incremental difference in the magnitude of an outcome, a categorical difference in delay has a greater impact in SE than in JE.

STUDY 6

Method

Study 6 was designed to test hypothesis 5 concerning risk preference involving a trade-off between certainty and magnitude. It examined respondents' willingness to fund two disaster prevention programs, one expected to protect a smaller number of people for sure (smaller riskless gain) and the other expected to protect a larger number of people with uncertainty (larger risky gain).

Participants were 81 students (33% female) from Shanghai Jiaotong University in China. In a questionnaire, they were asked to assume the role of a government official in charge of disaster prevention programs. They were then asked to imagine that a major flood was about to occur around the Yellow River and that tens of thousands of residents in the region would be affected. Two alternative prevention programs had been proposed. If one program was implemented, 20,000 residents would be protected from the flood. If the other program was implemented, there was a 90% chance that 40,000 residents would be protected and a 10% chance that none would be protected. Participants were randomly assigned to one JE condition or one of two SE conditions. Those in the JE condition were presented with both options (side by side, order counterbalanced), told that government funding was highly limited, and asked to indicate the maximum amount they would pay for each program by circling one of 10 options ranging from ¥10 million to ¥100 million. Respondents in each SE condition received the same instructions, except that they were presented with only one of the two programs and indicated their willingness to pay for that program.

Results and Discussion

According to hypothesis 5, relative to larger risky gains, smaller riskless gains will be favored more in SE than in JE. The results, summarized in table 2, were consistent with this prediction. A hybrid t-test revealed a significant JE-SE preference reversal ($t(78) = 2.38$, $p < .05$). In JE, respondents would pay more for the risky program ($M_{\text{risky}} = 70.67$ vs. $M_{\text{riskless}} = 58.33$, $t(29) = 2.03$, $p = .05$); in SE, re-

spondents would pay somewhat (not significantly) more for the riskless program ($M_{risky} = 70.00$ vs. $M_{riskless} = 79.13$, $t(49) = -1.22$, NS).

STUDY 7

Methods

Just as study 5 was a replication of study 4 with real consequences, so was study 7 a replication of study 6 with real consequences. Study 7 examined students' willingness to participate in two upcoming studies, one offering a smaller riskless payment and the other offering a larger risky payment.

Participants were 159 students (36% female) from Shanghai Jiaotong University in China. They were randomly assigned to either a JE condition or one of two SE conditions. The procedure of this study was identical to that of study 5 except for the payment of the upcoming studies. Here, one of the upcoming studies would pay each participant ¥20 for sure, and the other upcoming study offered each participant a 5/6 (83.3%) chance at ¥40, and the payment would be determined by rolling a die at the end of the study (¥40 if the outcome was 1–5 or ¥0 if the outcome was 6).

Results and Discussion

The results, summarized in table 2, replicated those of study 6. A hybrid *t*-test found a significant JE-SE preference reversal ($t(156) = 2.60$, $p = .01$). In JE, respondents were (marginally significantly) more willing to participate in the study that offered a larger risky payment ($M_{risky\ ¥40} = 2.69$ vs. $M_{riskless\ ¥20} = 2.47$, $t(51) = 1.78$, $p = .08$) whereas, in SE, respondents were more willing to participate in the study that offered a smaller riskless payment ($M_{risky\ ¥40} = 2.29$ vs. $M_{riskless\ ¥20} = 2.82$, $t(105) = -2.33$, $p < .05$). Together, studies 6 and 7 supported hypothesis 5, showing that, relative to the magnitude of an outcome, the certainty of the outcome exerts a greater influence in SE than in JE.

STUDY 8

Method

Study 8 was designed to test hypothesis 6 concerning trade-offs between incremental delay and categorical probability. It examined respondents' willingness to pay for two air conditioning repair services: one with a shorter delay and no guarantee of success and the other with a longer delay and guaranteed success.

Participants were 50 students (30% female) from the University of Chicago. In a questionnaire, they were asked to imagine that their air conditioning (AC) unit stopped working during a hot summer, that replacing it with a new one would cost $800, and that they were looking for an AC repair service to fix their unit at a lower cost. Of the two services, one could come in 3 days and had a 95% chance of fixing the AC unit, and the other would come in 12 days and had a 100% (guaranteed) chance of fixing the AC unit. For either service choice, a service fee must be paid up front and is nonrefundable even if the AC is not fixed. Participants were randomly assigned to either a JE condition or one of two SE conditions. Those in the JE condition were presented with both services and asked to indicate the maximum service fee they would be willing to pay for each service by indicating a number between $0 and $600. Those in each SE condition were told about only one of the services and indicated their willingness to pay for that service.

Results and Discussion

According to hypothesis 6, relative to less delayed risky gains, more delayed riskless gains will be favored more in SE than in JE. The results, summarized in table 2, supported this prediction. A hybrid *t*-test found a significant JE-SE preference reversal ($t(47) = 2.57$, $p < .05$). In JE, respondents would pay somewhat (not significantly) more for the less delayed risky service ($M_{less\ del.\ risky} = 126.47$ vs. $M_{more\ del.\ riskless} = 91.18$, $t(16) = 1.69$, NS) whereas, in SE, respondents would pay more for the more delayed riskless service ($M_{less\ del.\ risky} = 110.00$ vs. $M_{more\ del.\ riskless} = 187.50$, $t(31) = -2.04$, $p < .05$).

STUDY 9

Method

Study 9 was designed to test hypothesis 7 concerning trade-offs between incremental probability and categorical delay. It examined respondents' willingness to participate in a research study under two compensation plans: one with a smaller chance to receive a reward immediately and the other with a larger chance to receive the same reward with a delay.

Participants were 93 students (35% female) from the University of Miami. In a questionnaire, they were asked to imagine that a market research firm was recruiting college students to participate in a 3-hour study and that there were two compensation plans: one offering each participant an 85% chance at receiving $100 immediately afterward and the other offering each participant a 95% chance at receiving $100 1 month later. Participants were randomly assigned to a JE condition or one of two SE conditions. Those in the JE condition were presented with both compensation plans and asked to indicate how likely they would be to participate in the study under each plan by using a 7-point scale (1 = highly unlikely; 7 = highly likely). Those in each SE condition were told about only one of the two compensation plans and asked to indicate how likely they would be to participate in that study.

Results and Discussion

According to hypothesis 7, relative to higher probability delayed gains, lower probability immediate gains will be favored more in SE than in JE. The results, summarized in

table 2, confirmed this prediction. A hybrid t-test revealed a significant JE-SE preference reversal ($t(90) = 3.17$, $p < .01$). In JE, respondents were more willing to participate in the study that offered a higher probability delayed compensation ($M_{\text{hi prob. delayed}} = 5.97$ vs. $M_{\text{lo prob. immed.}} = 5.35$, $t(30) = 2.28$, $p < .05$) whereas, in SE, respondents were more willing to participate in the study that offered a lower probability immediate compensation ($M_{\text{hi prob. delayed}} = 4.10$ vs. $M_{\text{lo prob. immed.}} = 5.13$, $t(60) = -2.19$, $p < .05$). Together, studies 8 and 9 suggested that, like time preference and risk preference, preferences between time and risk themselves can also reverse between JE and SE when the difference in one attribute is categorical and the difference in the other is incremental.

GENERAL DISCUSSION

We have sought to make two contributions in this research. First, we have examined people's reactions to three basic attributes of choice options—magnitude, time, and probability—and connected these reactions to two basic modes of evaluation: JE and SE. Second, we have extended the analysis about single attributes to trade-off preferences involving two attributes, including two widely studied preferences—time preference and risk preference—and two less studied preferences, both about probability and time. In the remainder of this article, we explore implications of this research.

The Importance of Reference Points

According to our theory, the reason why the three classes of functions are more nonlinear around the endpoints in SE than in JE is that the endpoints serve as salient reference points, which define categories, and people in SE evaluate values by their categories. Theoretically, if there is a salient reference point other than the endpoints, the corresponding function will also be more nonlinear around that point in SE than in JE. To test this idea, we conducted another study in which we asked 209 students from Shanghai Jiaotong University in China to imagine that they just moved into their new home and were in a plant store shopping for an imported cactus ball to decorate their living room. We manipulated two variables: reference size (small vs. large) and evaluation mode (JE vs. SE). In the small-reference-size conditions, respondents were told that the average size of such imported cactus balls was 60 centimeters, and in the large-reference-size condition, the average size was 80 centimeters. In the JE conditions, respondents were told that the store carried three imported cactus balls with sizes 60, 70, and 80 centimeters, respectively. In each of the SE conditions, respondents were told about one of the three cactus balls. The dependent variable was willingness to pay. Notice that when the reference size was low (60 centimeters), the difference between 60 and 70 centimeters was categorical, and the difference between 70 and 80 centimeters was incremental; when the reference size was high (80 centimeters), the reverse was true. The results, summarized in table 1, were consistent with our theory. When the reference size was 60 centimeters, over the 60–70 centimeter range, respondents were similarly magnitude sensitive in JE and SE ($t(73) = .40$, NS), and over the 70–80 centimeters range, respondents were more magnitude sensitive in JE than in SE ($t(72) = 2.18$, $p < .05$). When the reference size was 80 centimeters, over the 60–70 centimeters range, respondents were more magnitude sensitive in JE than in SE ($t(76) = 3.12$, $p < .01$), and over the 70–80 centimeters range, respondents were only marginally more magnitude sensitive in JE than in SE ($t(76) = 1.86$, $p = .067$). (All the analyses used the hybrid t-test mentioned earlier.)

The study above illustrated one way of manipulating an ad hoc reference point. However, we believe that it is the endpoints, not such ad hoc values, that are the default and most salient reference points in most situations, and that the shapes of the three functions depicted in figure 1 are most common.

The Importance of Knowledge

Another implication from our theory concerns knowledge. The JE/SE differences in the three classes of functions as depicted in figure 1 arise only in situations where the target attribute, be it magnitude, time, or probability, is difficult to evaluate in isolation; these differences will not arise in situations where the target attribute is easy to evaluate in isolation (Hsee and Zhang 2010). Whether the target variable is easy or difficult to evaluate in isolation depends in part on the decision maker's knowledge about the variable in the given context. Relative to the average person, financial analysts are more knowledgeable about the risks of different investment options, and consequently their probability weighting function in SE (in which investment options are presented in isolation) will be more in line with their probability weighting function in JE (in which the investment options are presented together). We believe that it is the norm rather than an exception that the JE and SE functions diverge, because most decision makers are not experts.

Evaluations and Preferences in Real Life

In real life, sometimes we encounter multiple alternatives at the same time and are called upon to make a choice, as in JE, and sometimes we encounter one option at a time and need to decide whether to pursue it or not, as in SE. Our research indicates that judgments and choices made in these two modes can be quite different and suggests that failure to recognize the difference can lead to systematic and serious errors.

For example, when predicting consumers' sensitivity to the price of a new product, marketers often use comparative procedures, such as conjoint analyses, asking respondents to consider alternative prices and indicate their willingness to purchase, as in JE. However, in real life, consumers often encounter only one price for a given product—namely, the finally set price—and must decide whether or not to purchase the product at that price, as in SE. Consequently,

marketers may overestimate consumers' price sensitivity, namely, the steepness of their magnitude (price) valuation function over incrementally different prices. Marketers may predict that consumers would be much less willing to buy a new product if it were priced at $500 than at $450, but in reality, consumers may be as willing to buy it at $500 as at $450. Such prediction errors can cost marketers hefty profits. A recent paper has provided preliminary evidence for these propositions (Shen et al. 2012).

The same can be said of trade-off preferences. Risk preference and time preference are usually estimated in JE; respondents are asked questions such as "Do you prefer a $100 gift coupon for sure or an 80% chance at a $150 gift coupon?" and "Do you prefer a $100 gift coupon now or a $150 gift coupon in a month?" In real life, individuals often encounter such opportunities one at a time and need to decide whether to pursue the opportunities or not. Our research suggests that individuals may be more risk averse and more delay averse when encountering such opportunities in real life than when asked to express their preferences.

In this research, we have treated evaluation mode as if it were a dichotomous variable, either JE or SE. In reality, evaluation mode is a continuum. Evaluation mode refers to both how options are presented and how they are evaluated. How options are presented generally influences how they are evaluated. In the strongest JE case, the options are presented side by side so that they can be easily and directly compared. In the extreme SE case, each option is presented alone without any reference to the other option(s), so that they cannot be compared. In many real-world cases, the evaluation mode falls between the two extremes. For example, one option may be presented before another, or one option is in the foreground and another in the background. In both of these examples, the evaluation mode of the first option veers more in the SE direction than that of the second option (e.g., Moore 1999). Our research suggests that if the options involve trade-offs between magnitude and immediacy or between magnitude and certainty, the more the evaluation is in the SE direction, the more people will seek immediacy and certainty.

REFERENCES

Bartels, Daniel, and Lance J. Rips (2010), "Psychological Connectedness and Intertemporal Choice," *Journal of Experimental Psychology: General*, 139 (February), 49–69.

Bartels, Daniel, and Oleg Urminsky (2011), "On Intertemporal Selfishness: How the Perceived Instability of Identity Underlies Impatient Consumption," *Journal of Consumer Research*, 38 (June), 182–98.

Baucells, Manel, and Franz H. Heukamp (2012), "Probability and Time Trade-Off," *Management Science*, 58 (April), 831–42.

Bazerman, Max, Don A. Moore, Ann E. Tenbrunsel, Kimberly A. Wade-Benzoni, and Sally Blount (1999), "Explaining How Preferences Change across Joint versus Separate Evaluation," *Journal of Economic Behavior and Organization*, 39 (May), 41–58.

Bleichrodt, Han, and Jose L. P. Prades (2009), "New Evidence of Preference Reversals in Health Utility Measurement," *Health Economics*, 18 (September), 713–26.

Bommier, Antoine (2006), "Uncertain Lifetime and Intertemporal Choice: Risk Aversion as a Rationale for Time Discounting," *International Economic Review*, 47 (November), 1223–46.

Burson, Katherine A., Richard P. Larrick, and John G. Lynch Jr. (2009), "Six of One, Half Dozen of the Other: Expanding and Contracting Numerical Dimensions Produces Preference Reversals," *Psychological Science*, 20 (September), 1074–78.

Chatterjee, Subimal, Timothy B. Heath, and Junhond Min (2009), "The Susceptibility of Mental Accounting Principles to Evaluation Mode Effects," *Journal of Behavioral Decision Making*, 22 (July), 120–37.

Chen, Rong, and Jianmin Jia (2005), "Consumer Choices under Small Probabilities: Overweighting or Underweighting," *Marketing Letters*, 16 (March), 5–18.

Ebert, Jane E., and Drazen Prelec (2007), "The Fragility of Time: Time-Insensitivity and Valuation of the Near and Far Future," *Management Science*, 53 (September), 1423–38.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40 (June), 351–401.

Gneezy, Uri, John List, and George Wu (2006), "The Uncertainty Effect: When a Risky Prospect Is Valued Less than Its Worse Possible Outcome," *Quarterly Journal of Economics*, 121 (4), 1283–1309.

Gneezy, Uri, and Aldo Rustichini (2000), "Pay Enough, or Don't Pay at All," *Quarterly Journal of Economics*, 115 (August), 791–810.

Gonzalez, Richard, and George Wu (1999), "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38 (January), 129–66.

Hershfield, Hal E., Daniel G. Goldstein, William F. Sharpe, Jesse Fox, Leo Yeykelvis, Laura L. Carstensen, and Jeremy N. Bailenson (2011), "Increasing Saving Behavior through Age-Progressed Renderings of the Future Self," *Journal of Marketing Research*, 48 (November), S23–S37.

Hsee, Christopher K. (1996), "The Evaluability Hypothesis: An Explanation of Preference Reversals between Joint and Separate Evaluations of Alternatives," *Organizational Behavior and Human Choice Processes*, 46 (September), 247–57.

Hsee, Christopher K., and France Leclerc (1998), "Will Products Look More Attractive When Presented Separately or Together?" *Journal of Consumer Research*, 25 (September), 175–86.

Hsee, Christopher K., George Loewenstein, Sally Blount, and Max H. Bazerman (1999), "Preference Reversals between Joint and Separate Evaluations of Options: A Review and Theoretical Analysis," *Psychological Bulletin*, 125 (September), 576–91.

Hsee, Christopher K., and Yuval Rottenstreich (2004), "Music, Pandas and Muggers: On the Affective Psychology of Value," *Journal of Experimental Psychology: General*, 133 (March), 23–30.

Hsee, Christopher K., Fang Yu, Jiao Zhang, and Yan Zhang (2003), "Medium Maximization," *Journal of Consumer Research*, 30 (June), 1–14.

Hsee, Christopher K., and Jiao Zhang (2004), "Distinction Bias: Misprediction and Mischoice due to Joint Evaluation," *Journal of Personality and Social Psychology*, 86 (May), 680–95.

——— (2010), "General Evaluability Theory: An Analysis of When What We Care about Matters," *Perspectives on Psychological Science*, 5 (September), 343–55.

Kable, Joseph K., and Paul W. Glimcher (2010), "An 'As Soon

as Possible' Effect in Human Intertemporal Decision Making: Behavioral Evidence and Neural Mechanisms," *Journal of Neurophysiology*, 103 (5), 2513–31.

Kahneman, Daniel, and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47 (March), 263–91.

Kilka, Michael, and Martin Weber (2001), "What Determines the Shape of the Probability Weighting Function under Uncertainty?" *Management Science*, 47 (December), 1712–26.

Kluger, Avraham N., Elena Stephan, Yoav Ganzach, and Meirav Hershkovitz (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 (1), 20–39.

Kogut, Tehila, and Ilana Ritov (2005), "The Singularity Effect of Identified Victims in Separate and Joint Evaluations," *Organizational Behavior and Human Decision Processes*, 97 (April), 106–16.

Kunreuther, Howard, Nathan Novemsky, and Daniel Kahneman (2001), "Making Low Probabilities Useful," *Journal of Risk and Uncertainty*, 23 (September), 103–20.

Laibson, David (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112 (May), 443–77.

Laran, Juliano, and Keith Wilcox (2011), "Choice, Rejection, and Elaboration on Preference-Inconsistent Alternatives," *Journal of Consumer Research*, 38 (August), 229–41.

Lichtenstein, Sarah, and Paul Slovic (1971), "Reversals of Preference between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, 89 (1), 46–55.

List, John A. (2002), "Preference Reversals of a Different Kind: The 'More Is Less' Phenomenon," *American Economic Review*, 92 (December), 1636–43.

Loewenstein, George (1996), "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, 65 (March), 272–92.

Malkoc, Selin A., and Gal Zauberman (2006), "Deferring versus Expediting Consumption: The Effect of Outcome Concreteness on Sensitivity to Time Horizon," *Journal of Marketing Research*, 43 (November), 618–27.

Mischel, Walter, Yuichi Shoda, and Monica L. Rodriguez (1989), "Delay of Gratification in Children," *Science*, 244 (May), 933–38.

Moore, Don A. (1999), "Order Effects in Preference Judgments: Evidence for Context Dependence in the Generation of Preferences," *Organizational Behavior and Human Decision Processes*, 78 (May), 146–65.

Nowlis, Stephen M., and Itamar Simonson (1997), "Attribute-Task Compatibility as a Determinant of Consumer Preference Reversals," *Journal of Marketing Research*, 34 (May), 205–18.

Okada, Erica Mina (2005), "Justification Effects on Consumer Choice of Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 42 (February), 43–53.

Paharia, Neeru, Karim S. Kassam, Joshua D. Greene, and Max H. Bazerman (2009), "Dirty Work, Clean Hands: The Moral Psychology of Indirect Agency," *Organizational Behavior and Human Decision Processes*, 109 (July), 134–41.

Palmeira, Mauricio M. (2011), "The Zero-Comparison Effect," *Journal of Consumer Research*, 38 (June), 16–26.

Prelec, Drazen (1998), "The Probability Weighting Function," *Econometrica*, 66 (May), 497–527.

Prelec, Drazen, and George Loewenstein (1991), "Decision Making over Time and under Uncertainty: A Common Approach," *Management Science*, 37 (July), 770–86.

Rachlin, Howard, Jay Brown, and David Cross (2000), "Discounting in Judgments of Delay and Probability," *Journal of Behavioral Decision Making*, 13 (2), 145–59.

Read, Daniel (2001), "Is Time-Discounting Hyperbolic or Subadditive?" *Journal of Risk and Uncertainty*, 23 (July), 5–32.

Read, Daniel, Shane Frederick, Burcu Orsel, and Juwaria Rahman (2005), "Four Score and Seven Years from Now: The Data/Delay Effect in Temporal Discounting," *Management Science*, 51 (September), 1326–35.

Read, Daniel, and N. L. Read (2004), "Time Discounting over the Lifespan," *Organizational Behavior and Human Decision Processes*, 94 (May), 22–32.

Schmeltzer, Christopher, Jean-Paul Caverni, and Massimo Warglien (2004), "How Does Preference Reversal Appear and Disappear? Effects of the Evaluation Mode," *Journal of Behavioral Decision Making*, 17 (December), 395–408.

Scholten, Marc, and Daniel Read (2006), "Discounting by Intervals: A Generalized Model of Intertemporal Choice," *Management Science*, 52 (September), 1424–36.

Shamosh, Noah A., Colin G. DeYoung, Adam E. Green, Deidre L. Reis, Matthew R. Johnson, Andrew R. A. Conway, Randall W. Engle, Todd S. Braver, and Jeremy R. Gray (2008), "Individual Differences in Delay Discounting," *Psychological Science*, 19 (September), 904–11.

Shampanier, Kristina, Nina Mazar, and Dan Ariely (2007), "Zero as a Special Price: The True Value of Free Products," *Marketing Science*, 26 (November), 742–57.

Shen, Luxi, Christopher K. Hsee, Qingsheng Wu, and Claire I. Tsai (2012), "Overpredicting and Underprofiting in Pricing Decisions," *Journal of Behavioral Decision Making*, 25 (December), 512–21.

Slovic, Paul (1987), "Perception of Risk," *Science*, 236 (April), 280–85.

Tversky, Amos, Shmuel Sattath, and Paul Slovic (1988), "Contingent Weighting in Judgment and Choice," *Psychological Review*, 95 (July), 371–84.

Van den Bergh, Bram, Siegfried Dewitte, and Luk Warlop (2008), "Bikinis Instigate Generalized Impatience in Intertemporal Choice," *Journal of Consumer Research*, 35 (June), 85–97.

Weber, Bethany J., and Gretchen B. Chapman (2005), "Playing for Peanuts: Why Is Risk Seeking More Common for Low-Stakes Gambles?" *Organizational Behavior and Human Decision Processes*, 97 (May), 31–46.

Zauberman, Gal, B. K. Kim, Selin A. Malkoc, and James R. Bettman (2009), "Discounting Time and Time Discounting: Subjective Time Perception and Intertemporal Preferences," *Journal of Marketing Research*, 46 (August), 543–56.

Zauberman, Gal, and John G. Lynch Jr. (2005), "Resource Slack and Propensity to Discount Delayed Investments of Time versus Money," *Journal of Experimental Psychology: General*, 134 (October), 23–37.