

The Dissociation Between Monetary Assessment and Predicted Utility

On Amir

Rady School of Management, University of California at San Diego, La Jolla, California 92037,
oamir@ucsd.edu

Dan Ariely

Fuqua School of Business, Duke University, Durham, North Carolina 27706,
ariely@mit.edu

Ziv Carmon

INSEAD, Asia Campus, 138676, Singapore, ziv.carmon@insead.edu

We study the dissociation between two common measures of value—monetary assessment of purchase options versus the predicted utility associated with owning or consuming those options, a disparity that is reflected in well-known judgment anomalies and that is important for interpreting market research data. We propose that a significant cause of this dissociation is the difference in how these two types of evaluations are formed—each is informed by different types of information. Thus, dissociation between these two types of measures should not be interpreted as failure to map utility onto money, as such mapping is not really attempted. We suggest that monetary assessment tends to focus on the transaction in which the purchase alternative would be acquired or forgone (e.g., how fair the transaction seems), failing to adequately reflect the purchase alternative itself (e.g., the expected pleasure of owning or consuming it), which is what informs predicted utility judgments. We illustrate the value of this idea by deriving and testing empirical predictions of disparities in the impact of different types of information and manipulations on the two types of value assessment.

Key words: decision-making; buyer behavior; market research

History: This paper was received November 8, 2005, and was with the authors 9 months for 2 revisions; processed by Ravi Dhar. Published online in *Articles in Advance* September 16, 2008.

Introduction

Imagine finding out that later today your favorite singer was appearing in a one-time performance at a local venue that seats up to 500 people. Consider your reservation price for a ticket to this concert, i.e., the price beyond which you would refuse to purchase the ticket, thus forgoing the opportunity to attend the event. Would it matter if the fee that the producer was asked to pay for the show by the local municipality amounted to \$2,500 or \$45,000? Would it matter whether you have a sore muscle? Fees paid by the producer would have little or no impact on the pleasure you would derive from attending the concert, yet they would significantly affect the reservation price. Physical discomfort experienced during the concert would certainly influence the pleasure you could derive from attending yet it would have little impact on your reservation price.

This example illustrates a disparity between willingness to pay and predicted experience. Dissociation between monetary assessment and predicted utility is troublesome (even though it may be accounted for by

market considerations; cf. Kreps 1990). That is because measures such as reservation prices express the desire to engage in a transaction, presumably reflecting the utility associated with what is to be acquired or forgone in that transaction. In other words, the dissociation conflicts with the assumption that reservation price and predicted utility reflect the same construct—utility. This is important to marketing researchers. Conjoint analyses, for example, assume that differences in price points can be readily translated into utility differences. Reservation price assessment is commonly used to measure the desirability of products, services, and public policy initiatives. But, as the concert ticket example illustrates, monetary assessment can be dissociated from the utility that consumers expect to derive from the evaluated alternative. Such disparities between the two types of assessment may also underlie well-known judgment anomalies such as the Money Illusion (Shafir et al. 1997), decision inconsistencies (Hsee 1999), and Coherent Arbitrariness (Ariely et al. 2003), as well as fairness effects (Kahneman et al. 1986, Thaler 1985), presenting both theoretical and practical difficulties.

We propose and test an account for the dissociation, whereby predicted utility and monetary assessments are formed differently, each informed by different types of variables. Thus, the dissociation is not necessarily caused by a failure to map utility onto the two measures. Specifically, monetary assessments such as reservation prices focus on variables relating to the transaction in which the purchase alternative would be acquired or forgone (e.g., how fair it seems). As a result, monetary assessments tend to reflect factors that do not significantly influence the pleasure associated with owning or consuming the purchase alternative (such as its production cost), failing to sufficiently reflect other factors that do influence pleasure (such as how pretty the item is).

To support the validity of our account and to illustrate its value, we derive and test several empirical predictions. In Experiment 1, we demonstrate the disparity. In Experiment 2, we show that drawing attention to the consumption experience (enabled by the transaction) makes monetary evaluations more consistent with corresponding assessments of predicted utility. In Experiment 3, we further demonstrate the sensitivity of monetary evaluations to transaction-related variables, rather than to variables relating to the experience (provided by the item that would be acquired). In Experiment 4, we demonstrate another difference between the two types of measures, i.e., although experience-related variables have an impact on monetary evaluations if they are made very salient, monetary evaluations have little impact on predicted utility even if they are made salient.

In the next section we present our account and empirical predictions that follow from it. We then describe four main studies as well as several other empirical tests. Finally, we summarize our research and its findings and discuss implications and directions for future research.

How Monetary and Pleasure Assessments Are Different

Dissociation between monetary assessment and predicted utility (such as that reflected in the example opening this paper) underlies a variety of well-known preference anomalies. One example is Thaler's (1985) "beer on the beach" scenario, in which individuals were asked to state their reservation price for beer that would be consumed on the beach but purchased by a friend either from a fancy hotel or from a run-down local grocery store. Thaler found that people were willing to pay considerably more for beer bought at a fancy hotel, although the pleasure of consuming the beer would presumably not be influenced by where the beer was purchased. In another example, Hsee et al. (1999) showed that people

chose options that were valued more according to an imaginary currency, even when this conflicted with their predictions of which option would yield greater utility. In yet another example, Ariely et al. (2003) first asked people if they would purchase various products at a price made of two digits from their social security number (e.g., 34 became \$34), then elicited reservation prices for each product. Although the two digits were obviously not predictive of product utility, they greatly affected reservation prices, presumably because they were a significant transaction cue—determining the first price considered by the people questioned.

In examples such as the one presented at the beginning of this paper, price assessments focus on features of the potential transaction in which the purchase would be made or forgone, and insufficiently reflect the predicted utility associated with that purchase alternative. This disparity can be traced to focalism, i.e., the notion that different assessment tasks tend to be naturally informed by different features of the evaluated stimuli. For example, Temporal Construal Theory (Trope and Liberman 2003) suggests that assessments made for the near future are informed more by the feasibility of accomplishing a goal, whereas long-term assessments are informed more by the desirability of the outcomes. In another example, Wilson et al. (2000) and Schkade and Kahneman (1998) show that people tend to overestimate the impact of events that seem significant to them (e.g., a defeat by their favored sports team, or relocating to a state with better weather) on their overall well being. This is because well-being judgments are too heavily informed by this salient event, thus neglecting the impact of other aspects of those people's lives. In another example, the Prominence Effect (Tversky et al. 1988) traces the cause of preference reversals, such as disparity between choices and ratings, to choices that are heavily informed by the most prominent attribute, whereas ratings are informed by less prominent attributes of the evaluated options. Finally, Shiv and Huber (2000) find that if allowed, consumers tend to focus on available mental imagery when predicting satisfaction more than when choosing. They show that when consumers consider purchase satisfaction they use decision strategies that emphasize and devote more attentional resources on vivid attributes, thus generating different choices. Similarly, our account suggests that people naturally focus on different cues and attributes of the decision when forming monetary assessments versus utility predictions. We conceptualize experience cues as including cues that enable mental imagery, but also any cues that make predicting the experience easier. Our account is in line with Shiv and Huber (2000) with regard to explaining utility predictions, though

our focus is not on comparison to choice, but rather to monetary assessments, a seemingly closer form of judgment.

Building on the notion of focalism, we suggest that monetary assessment and predicted utility, the two types of assessment we explore in this paper, are formed differently, and that each tends to be informed by different types of information. Thus, dissociation between the two types of assessment does not necessarily reflect poor mapping of utility onto monetary assessments. Rather, monetary assessments may not naturally reflect utility as they focus on the transaction. More specifically, monetary assessments naturally focus outward on what we term *transaction cues*, i.e., variables related to assessment of the transaction such as costs, reference prices, and market norms (cf. Kahneman et al. 1990, Winer 1988). Utility predictions, on the other hand, focus inward on what we term *experience cues*, i.e., aspects of the experience of owning or consuming the purchase alternative. Thus, consumers estimate the utility associated with owning or consuming the purchase alternative based on mental representations such as experiences and schemas (cf. Kahneman and Miller 1986, Shiv and Huber 2000).

This difference in how the two assessments are formed allows us to make empirical predictions. The basic claim (first tested in Experiment 1) is that *monetary evaluations naturally focus on transaction cues, while utility predictions naturally focus on experience cues*. A second claim (tested in Experiment 2) based on this idea is that focusing on experience cues will have a greater impact on monetary evaluations, which do not naturally focus on such cues, than on predicted utility judgments that naturally do. Moreover, providing transaction cues, but not experience cues, will have significant impact on monetary evaluations as these are the cues that naturally feed into these evaluations. Because monetary evaluations rely heavily on transaction cues, without such cues these evaluations will be diffuse and imprecise (this is tested in Experiment 3). Finally, we demonstrate (in Experiment 4) another manifestation of the difference in how the two types of judgments are formed—an asymmetry stemming from a difference between transaction and experience cues. Availability of predicted utility affects monetary evaluations, but availability of monetary assessments has little or no impact on utility predictions. Thus, when decision makers form monetary judgments they pay little attention to experience cues because of focalism, rather than because they believe that such cues are irrelevant to monetary assessments. The opposite, however, is not true—Decision makers do not naturally find transaction cues relevant to predicted utility.

Experiment 1: Differential Impact of Transaction- and Experience-Related Variables on Monetary vs. Pleasure Assessments

In our first experiment we test our basic prediction about the differential impact of transaction- and experience-cues on monetary versus pleasure assessments.

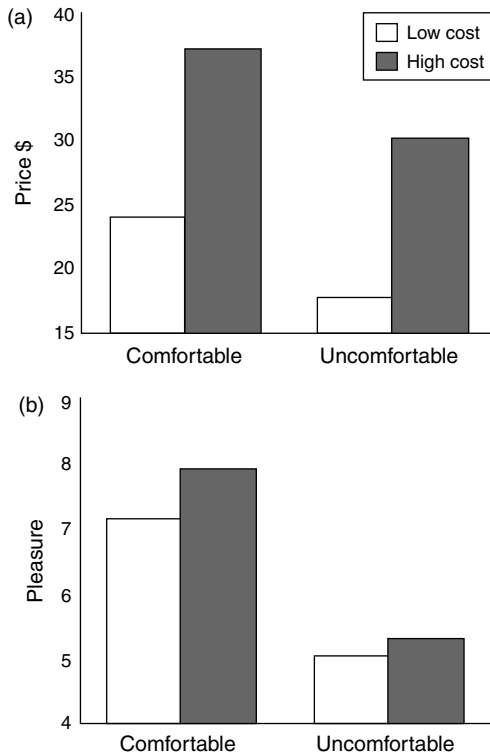
Method

Three hundred and eighty-eight respondents participated in the study by completing a brief scenario-based survey. The scenario described a concert of a popular singer planned by a student association. The transaction cue that we manipulated, i.e., the cost of the event to the student association, was described as either \$12,500 (low cost) or \$45,000 (high cost). The experience cue that we manipulated, i.e., the expected temperature in the auditorium, was to be either 70 degrees Fahrenheit (comfortable), or 95 degrees Fahrenheit (uncomfortable) depending on whether the air-conditioning system would operate. Half of the respondents were asked for their reservation price—their maximum willingness to pay (WTP) for a ticket to the concert by presenting them with a series of 20 choice questions, querying whether or not they would agree to pay a sum of \$X for the ticket, X ranging from \$6 to \$63. The other half of the respondents were asked to indicate their anticipated pleasure from this performance on a scale from 1 (not pleasurable) to 10 (very pleasurable). Thus, the study consisted of a 2 (cost of production: high versus low) × 2 (temperature at the auditorium: comfortable versus uncomfortable) × 2 (response type: reservation price versus predicted pleasure) between participants design.

Results and Discussion

We first analyzed the monetary responses in a cost of production (2) × temperature at the auditorium (2) analysis of variance (ANOVA), finding a main effect for production cost [$F(1, 188) = 64.11, p < 0.001$], and a main effect for temperature [$F(1, 188) = 19.39, p < 0.001$]. As shown in Figure 1(a), production cost had a greater impact on reservation prices than temperature changes (effect sizes of 0.51 and 0.28, respectively¹). Next, we analyzed the pleasure responses in a production cost (2) × temperature at the auditorium (2) ANOVA, finding a main effect for production cost [$F(1, 192) = 4.96, p = 0.027$] and a main effect for temperature [$F(1, 192) = 119.71, p < 0.001$]. As shown in Figure 1(b), production cost had much less impact on expected pleasure than temperature changes (effect

¹ Effect sizes are calculated as Cohen's *d*.

Figure 1 Reservation Prices and Predicted Pleasure Ratings for the Concert

sizes of 0.025 and 1.675, respectively). While it may seem questionable whether production cost should have any effect on the pleasure participants expect to derive from the concert, our respondents may have implicitly or explicitly assumed that a more expensive production could be more pleasurable (cf. Shiv et al. 2005). Regardless, analyses indicate that production cost had a greater impact on monetary responses, while the auditorium temperature had more impact on anticipated pleasure, as predicted.

Because the response scales for the two dependent measures were different, we next converted responses to these two measures to Z-scores. Analysis of the Z-scores in production cost (2) \times temperature at the auditorium (2) \times response scale (2) ANOVA revealed main effects for production cost [$F(1, 380) = 54.66$, $p < 0.001$] and for temperature [$F(1, 380) = 113.92$, $p < 0.001$]. More important, this analysis revealed an interaction of production cost with response type [$F(1, 380) = 19.04$, $p < 0.001$] and an interaction of temperature with response type [$F(1, 380) = 17.71$, $p < 0.001$]. This Z-score analysis supports our conclusion that monetary responses were more sensitive to the transaction cue, while pleasure responses were more sensitive to the experience cue.

In sum, Experiment 1 illustrates that monetary responses (reservation prices) were informed more by production cost (a transaction cue) than by expected

temperature (an experience cue), whereas the opposite was true for predicted utility (enjoyment ratings). To further support our account of the difference between monetary evaluations and predicted utility, in the next study we demonstrate that greater attention to experience cues moderates the effect of transaction cues on monetary assessments but has little impact on predicted utility.

Experiment 2: Differential Impact of a Focalism Manipulation

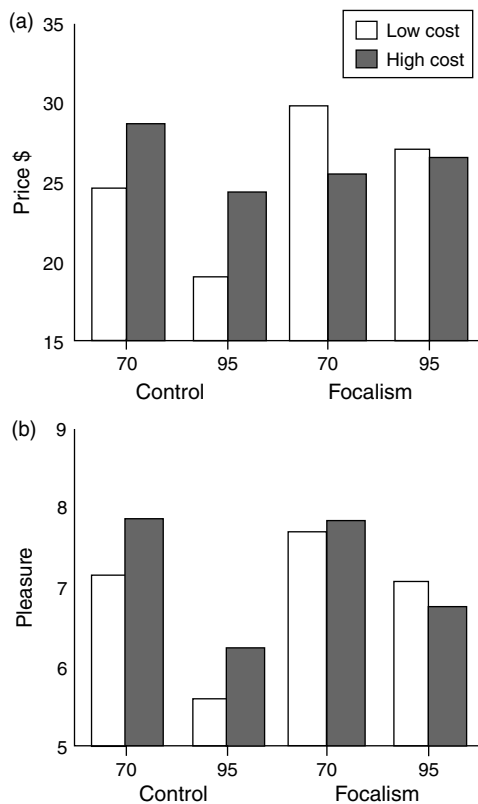
Method

Four hundred and forty-seven students at a major West Coast university participated in this survey that was run right before the student council annual concert. We used the same scenario as in Experiment 1. The study consisted of a 2 (response type: predicted pleasure versus reservation price) \times 2 (focalism: focusing questions versus none) \times 2 (cost of production: \$50,000—high versus \$5,000—low) \times 2 (temperature in the auditorium: comfortable versus uncomfortable) between participants design. Thus, for half of its participants, Experiment 2 replicated Experiment 1. The other half saw the same concert description, but before we administered the main dependent variables we asked three questions to draw participants' attention to the concert experience. Participants were asked to describe: "...how comfortable you predict the seats in the auditorium would be," "...who you might go with to this concert," and "...where you would hope to sit relative to the stage." As in Experiment 1, the dependent measures of interest were the reservation price and predicted pleasure rating. Reservation prices were assessed via a series of choice questions on whether the participant would agree to pay for the ticket each of 11 prices ranging from \$0 to \$50 in steps of \$5. Predicted pleasure was measured with a scale ranging from 0 (really not enjoy) to 7 (really enjoy).

Results and Discussion

As predicted, the results of the conditions that replicated Experiment 1 showed the same effects: Relative to their counterparts who indicated monetary evaluations, participants in the pleasure prediction conditions were affected more by the temperature during the concert than by the cost of production (Figure 2) [effect sizes of 0.415 and 0.054 versus 0.036 and 0.032, respectively]. Thus, participants were not willing to pay a higher price for the more enjoyable concert (uncomfortable temperature but costly to produce versus the comfortable temperature but less costly to produce). More important, participants who answered the focalism questions responded differently than those who did not: Focalism had marginal

Figure 2 Average WTP and Predicted Pleasure for the Concert With and Without a Focalism Manipulation



impact on the effect of cost on pleasure ratings [$F(1, 224) = 2.86, p = 0.092, \lambda = 2.86$], but a large impact on the effect of cost on monetary evaluations [$F(1, 223) = 3.59, p = 0.05, \lambda = 3.59$]. In fact, the latter impact was so pronounced that under the focalism manipulation, there was no longer a significant effect of cost on WTP [$F(1, 111) = 0.81, ns.$]. After focusing attention on the experience, participants virtually ignored the production cost information, the same information that previously dominated their monetary responses.

Apart from replicating the results of Experiment 1, Experiment 2 provides additional support for our account. Consistent with our proposition about the different focus of the two assessment types, we show that making experience cues more salient significantly reduced the impact of cost information (a transaction cue) on monetary evaluations. This also supports our notion that when people form monetary evaluations they naturally focus on transaction cues rather than on experience cues that may reflect the utility they would derive from the purchase alternative. We also demonstrate that the impact of transaction cues can be mitigated by drawing attention to experience cues.

Thus far, our studies demonstrate that consumers' reservation price assessments can be dissociated from predicted enjoyment. Monetary judgments appear to

focus on transaction cues whereas predicted pleasure judgments rely on experience cues. In Experiment 3 we test another prediction derived from our account. Specifically, if monetary assessments focus on transaction (rather than experience) cues, then in the absence of transaction cues, monetary assessment will be difficult to generate and thus be diffuse across respondents. Moreover, added experience cues will not make monetary evaluations easier or decrease this dispersion. But as monetary assessments naturally focus on transaction cues, adding such cues will simplify monetary evaluations and decrease their variability.

Experiment 3: Differential Impact of Transaction vs. Experience Cues on Monetary vs. Pleasure Assessments

In this experiment we examined the different impact of transaction versus experience cues on the two types of assessment in the domain of compensation—the lowest sum for which participants would agree to complete a task (i.e., willingness to accept—WTA). We used two types of tasks: some with which our participants were more familiar and others with which they were less familiar. More specifically, we pretested different tasks for familiarity and perceived effort, and chose tasks that did not differ in the distribution of the latter (more on this below). Among those tasks we selected two with which participants indicated that they were very familiar, and two with which they said they were very unfamiliar. We refer to such tasks as *market* and *nonmarket tasks*, respectively. As task familiarity correlates with knowledge of rough compensation paid for the task, we expected market tasks to be associated with significantly stronger transaction cues (standards of comparison) relative to nonmarket tasks.

Building on our argument that transaction cues inform monetary assessments, we expected scarcity of such cues to result in high variance across participants. Thus, we predicted that for market tasks variance for the minimal compensation demanded would be low, as individual estimates would anchor on prevailing rates and thus be relatively similar across individuals. On the other hand, because there would be less salient prevailing rates for nonmarket-tasks we expected high variance for such tasks. Importantly, as we argue that monetary assessment is mostly informed by transaction cues, we expected that providing experience cues (details on the subtasks and the effort required) will not significantly influence monetary assessments.

To test our predictions we manipulated the presence of two types of information. The first was experience cues in the form of detailed descriptions of the tasks, enhancing predictability of the experience

(and hence the utility associated with it). The second type of information was transaction related, in the form of the typical wage for a related task. We predicted that the first type of information would not influence the variance in monetary (compensation) responses, whereas the second type of information would decrease variance for nonmarket tasks (i.e., tasks that do not evoke a reference wage).

Method

One hundred and fourteen individuals participated in an online survey in return for a chance to win a free DVD. The study consisted of a 2 (information provided about the task: detailed versus none) \times 2 (reference wage information: provided versus not) \times 2 (task type: market versus nonmarket) mixed design, with the first two factors manipulated between participants, and the latter within participants. Based on a pretest we chose a set of tasks that did not differ in the distribution of perceived effort (see Table 1). The conditions differed in the amount and type of information provided to participants about each task. Half of the participants received a detailed description of the task, while the other half did not. For example, the description of the task of making coffee for three hours was supplemented by—This task is comprised of making cups of coffee next to the student center; the coffee should, as much as possible, be made in a standard fashion. Crossed with this manipulation, half of the participants received information about a related reference wage (“In a [related task] professionals earn \$x.xx per hour”), while the other half did not. For example, the task of making coffee for three hours was supplemented by “People who work at local shops usually earn around \$15 an hour.”

After reading the instructions, participants were asked for the minimal wage for which they would agree to perform the task (i.e., their WTA). We emphasized that this amount meant that they would not agree to perform the task if they were offered any less than the sum they stated. After indicating their WTA participants rated the effort each task required (as a measure of disutility), and their familiarity with the task (as a manipulation check).

Results and Discussion

The familiarity manipulation check relied on an eleven-point scale (ranging from -5 to $+5$), and

it revealed that the classifications of tasks had the intended effect. While for the market tasks (preparing coffee and helping elderly people) the mean familiarity ratings were 2.17 and 3.72, respectively, the means for the nonmarket tasks (shining shoes and preparing burgers) were -2.04 and -3.75 [smallest $t(55) = 3.65$, $p < 0.001$].

Recall that our main prediction focused on the impact of the different types of information on the variance in WTA assessments. The main effect of the market versus the nonmarket tasks was as predicted (see Table 1): Variance of WTA for market tasks was significantly smaller than that for nonmarket tasks [F -values ranged from $F(113, 113) = 8.69$ to $F(113, 113) = 33.68$, $p < 0.001$].² Beyond this main effect, we examined what it was about market tasks that substantially reduced the variance³—whether it was information allowing better prediction of the hedonic impact of the task or information about related wages. Our theory predicts the latter rather than the former, and as can be seen in the following analysis, the results support this: The effect of task specific information (experience cues) on the WTA variance was small, whereas the effect of related market wages (transaction cues) was large.

The following pattern emerges from analyzing the effects of the two types of information on the variance in WTA responses for market and nonmarket tasks: Compared to the no added information condition, adding reference wage information had greater impact on responses to the nonmarket tasks than on market tasks ($F[113, 113] = 10.34$ versus 2.32), and a greater effect than task information ($F[113, 113] = 2.26$ versus 1.97) and both types of information combined ($F[113, 113] = 3.37$ versus 1.93). Moreover, note that the relative effect of reference wage information on nonmarket tasks was larger than that of any other effect. It is also important to demonstrate that the variance in WTA does not stem from variance in expectations about the effort required for the different tasks (Table 1). Indeed, as can be seen in Figure 3 and consistent with our predictions, variance in expected effort (i.e., disutility) for the two types of tasks across the four information conditions was statistically equivalent (largest $F = 1.24$, *ns.*).

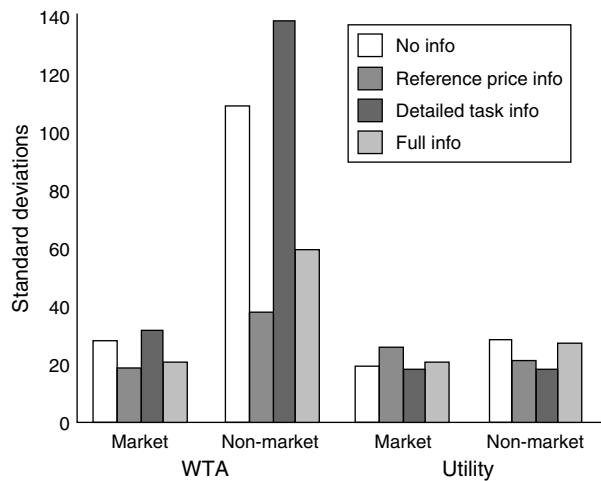
Table 1 Standard Deviations of WTA for Familiar and Unfamiliar Tasks Without Added Information

Task type	Task (all 3 hours long)	Std. dev. WTA	Std. dev. effort
Market	Prepare coffee	37.50	4.61
	Help elderly people	19.14	4.62
Nonmarket	Shine shoes	110.55	5.04
	Prepare burgers	111.08	5.62

² The difference in the variance could be assessed by their ratio, which provides an F -statistic, and because our sample sizes are identical we can directly compare the magnitude of this statistic as a proxy for effect size.

³ An alternative account for which we found little support was that differences in the variances related to differences in the mean WTA levels. Projecting the difference in the variances that may be caused by mean shifts, we find that those may cause the variance to change by up to 10%, two orders of magnitude less than the differences we observe.

Figure 3 Standard Deviations of WTA and Utility Predictions for Market and Nonmarket Tasks, With and Without Detailed Task and Reference Price Information



These results support our prediction that familiarity with market wage for a task leads to a significantly narrower distribution of WTA responses. Experiment 3 further supports this by demonstrating that an external reference wage substantially decreased variance, while task-specific information which can improve the internal representation of the effort involved did not (see Figure 3). Furthermore, consistent with our proposed account, utility predictions were virtually not influenced by such transaction cues. The results reinforce the notion that individuals do not naturally use internal representations of hedonic aspects of experiences to generate monetary assessments (hence the high variance). Rather they tend to rely on external transaction cues for their monetary assessments.

Experiment 4: Monetary Assessment Can Better Reflect Predicted Utility

The first three experiments consistently support our claim that monetary assessments emphasize transaction cues versus experience cues. Our final experiment examined another difference between monetary assessments and predicted utility: If monetary assessments insufficiently incorporate predicted utility because they naturally rely on transaction cues, explicitly (artificially) evoking predicted utility should change that. Thus, if people are first asked to assess predicted utility, this will influence their subsequent monetary assessments. Note that we are assuming that people recognize, implicitly or explicitly, that it is appropriate for monetary assessments to reflect predicted utility. In other words, we are assuming that people fail to naturally incorporate predicted utility information into monetary assessments even though they find merit in using such

information. In contrast, as we expect that people will not find merit in their own monetary assessments as information that should affect predicted utility,⁴ we expect predicted utility not to be influenced by whether monetary assessments are first elicited. Demonstrating such asymmetry would further support our proposition that the two types of assessments are formed differently.

Experiment 4 assessed the willingness to pay a technician who would recover lost data from a defective hard disk. Whereas in Experiment 3 the transaction cue we provided was the typical wage for a related activity, in Experiment 4 we offer information about the amount of labor the evaluated service would require. Consistent with our basic premise, we predicted that maximal WTP would be informed by the amount of labor required (transaction cue), more than by the importance of the service (i.e., experience cue reflecting predicted utility). Moreover, if predicted importance were first explicitly evoked (and thus made cognitively available) then it would inform WTP. On the other hand, making the amount of labor explicit would not influence assessment of the expected value of the service because this type of information does not seem relevant for this type of assessment.

Method

Ninety-six respondents completed a brief survey. The basic design was similar to that of the previous experiments, with the addition of an order factor: 2 (duration of work needed to repair the drive: 5 minutes versus 12 hours) × 2 (loss-magnitude: 1 month of data versus 5 years of data) × 2 (response mode order: first the WTP versus first the importance of the data recovery) between participant design. Participants were asked to imagine that they had just experienced a hard disk crash, and lost 1 month [5 years] of data, and that the technician who could restore their data would need 5 minutes [12 hours] to complete the task. Two questions were asked: the highest sum (reservation price) that participants would agree to pay for the data recovery, and rating of the importance participants assigned to the data recovery (importance, reflecting the displeasure or the disutility associated with not retrieving the data), on a scale of 0 (Not important) to 100 (Extremely important).

Results and Discussion

The results (see Table 2) replicated those of previous experiments: The labor required to restore the data had a main effect on WTP, but only the magnitude of the loss affected importance. This resulted

⁴ For example, that the reference price of a sandwich is greater than that of an ice cream cone should not imply that the former provides greater utility.

Table 2 Averages of WTP and Importance Judgments

Work length	Data lost	WTP first		Importance first	
		WTP (\$)	Importance	WTP (\$)	Importance
5 minutes	1 month	62.31	54.62	90.25	69.17
	5 years	127.50	71.75	114.50	85.50
12 hours	1 month	154.58	68.50	91.67	65.33
	5 years	213.46	74.08	184.08	86.25

in the paradoxical result that WTP for 5 minutes of work restoring 5 years of data was lower than the WTP for 12 hours work restoring 1 month of data. Furthermore, when importance ratings were elicited first, subsequent WTP responses corresponded to it and the paradoxical result was eliminated. Thus, WTP assessments more closely reflected predicted utility when the latter was made salient and accessible. Otherwise, consistent with the results of our previous studies, monetary responses more closely mirrored the amount of labor required.

These observations are supported by formal analysis. An ANOVA of the WTP responses elicited before the importance question was asked, including labor-duration and loss-magnitude as independent factors reveals a main effect for labor duration [$F(1, 46) = 6.24, p = 0.016$], but only a marginal effect for the loss magnitude [$F(1, 46) = 3.02, p = 0.09$]. When WTP was elicited after the importance rating was provided, however, we find the reverse—we do not find a significant main effect for labor duration [$F(1, 42) = 1.69, p = 0.201$], but rather a significant main effect for the loss magnitude [$F(1, 42) = 4.94, p = 0.031$]. Thus, respondents who first considered the importance of the repair used different information to form their WTP responses: They focused on the magnitude of the data loss only after being asked about it. Otherwise, they were more influenced by the amount of labor required to repair the drive.

Also as predicted, when judging the importance of recovering the data, participants were only influenced by loss magnitude. They were not influenced by the amount of labor required to restore the data regardless of whether they answered this question before or after they reported their WTP. Indeed, an ANOVA of the importance ratings reveals only a main effect for the amount of data lost when importance was elicited first and when it was elicited second [$F(1, 42) = 6.31, p = 0.016$; and $F(1, 46) = 3.81, p = 0.057$, respectively]. We find no effect for the expected labor duration in both conditions [$F(1, 42) = 0.43, ns.$; and $F(1, 46) = 0.96, ns.$, respectively]. Importantly, this refutes a potential alternative explanation whereby the impact of first eliciting importance ratings affected subsequent monetary evaluations because of demand effects.

General Discussion

The dissociation between two types of valuation, monetary assessment and predicted utility, is reflected in well-known preference anomalies (e.g., Kahneman et al. 1990, Shafir et al. 1997, Simonsohn and Loewenstein 2006, Sunstein et al. 2002, Thaler 1985), presenting practical and theoretical challenges. In this paper we propose that a significant cause of this dissociation is that the two types of assessment are formed differently, as each tends to be informed by different types of information. Monetary assessments heavily reflect transaction cues (i.e., variables relating to the transaction in which the purchase alternative may be acquired or forgone), insufficiently reflecting experience cues (i.e., variables relating to the experience of owning or consuming the purchase alternative) that inform predicted utility. Thus, the dissociation between the two types of assessment is not simply caused by poor mapping of utility onto monetary assessments as people may not attempt such mapping (cf. Carmon and Ariely 2000). Instead, they base assessments on considerations that naturally come to mind for the particular type of judgment.

In four studies we find consistent support for our basic proposition and for predictions we derive from it. In Experiment 1 we demonstrate dissociation between monetary assessment and predicted utility, and propose that this can be caused by a difference in how the two judgments are formed—different sources of information on which each type of judgment focuses. Specifically, we find that the expected temperature in the auditorium in which a concert is to take place, an experience cue, influenced the expected pleasure from attending a concert more than it affected the reservation price for a ticket to this concert. Conversely, the cost of producing the concert, a transaction cue, affected the reservation price more than expected pleasure. In Experiment 2 we demonstrate that drawing respondents' attention to experience cues (e.g., asking them to consider how comfortable their seat would be) increased the impact of experience cues on monetary assessments relative to their impact in a control condition. This led reservation prices to be more closely associated with predicted utility. Otherwise, without this forced focus, the monetary assessments seemed to naturally focus on transaction—rather than experience—cues.

Another prediction we derived from our notion of how the two types of valuation are formed was that monetary assessments will be diffuse absent salient transaction cues (which monetary assessments naturally depend on) even in the presence of significant experience cues. In Experiment 3 we find support for this prediction, in the lowest compensation level (i.e., WTA) at which participants indicated that they were

willing to perform different tasks. Some of those tasks evoked a reference wage more readily than others. We manipulated explicit availability of experience cues (description of what the task entails) and of transaction cues (prevailing wage for a similar task). As predicted, we found that WTA for performing tasks with less salient cues of common compensation (a transaction cue) varied considerably across respondents compared to WTA for tasks with more salient cues. Moreover, monetary assessments, but not predicted utility judgments, were influenced by the provision of transaction cues (common compensation for a related task) but not by the provision of experience cues (a detailed description of what the tasks entailed).

Finally, in Experiment 4 we illustrated another difference between the two types of judgments: Monetary assessments reflected predicted utility more closely when the former judgments were elicited before the latter, whereas predicted utility was not influenced by whether it was elicited before or after the monetary assessments. We attribute this to people's sense that experience cues (such as how enjoyable it will be to consume the purchase alternative) should be reflected in monetary assessments, but transaction cues (such as common wages) need not be significantly reflected in utility predictions. In other words, the results of this study suggest that the dissociation between the two types of measures is not caused by beliefs about what factors should influence monetary judgments. Rather, it is because predicted utility is not naturally salient when monetary judgments are formed. Specifically, in Experiment 4, WTP for restoring data from a crashed hard disk did not reflect predicted utility, but this dissociation was significantly reduced when predicted utility was elicited before the corresponding monetary evaluations. The study further illuminates a simple approach to reducing the dissociation between the two types of valuation: Monetary judgments can better reflect predicted utility if the latter assessments are elicited before the former.

In addition to describing how the two processes differ, our framework allows us to conjecture when these two types of assessments will be associated more or less closely. Given that the two types of evaluations tend to be informed by different types of cues, the evaluations will be consistent *if* the two types of cues are aligned. Thus, when market-related (transaction) cues correspond to one's utility function, or when the individual is made to explicitly predict utility before the monetary evaluation, we expect to find less dissociation of the type we investigated. For example, as customers consider their reservation price for a new house a realtor could encourage them to imagine what it would be like to live in that house to

help them rely less heavily on variables of questionable relevance (but shown to have significant impact; Simonsohn and Loewenstein 2006) such as the rent they paid for their previous apartment.

Implications of these ideas go beyond consumer judgments and market research. In labor markets for example, job seekers often know a particular salary level, such as the salary of a colleague. Such information (transaction cue) may cause people not to accept jobs that may otherwise greatly appeal to them because of heavy focus on comparative pay rather than (as long as the pay is high enough to cover their needs), on how much pleasure or pride they can expect to derive from the job. Thus, potential employees may demand the salaries they *deserve* based on market variables and thus fail to accept jobs that could bring them great satisfaction. On a related note, our results suggest that the tendency to make salaries in various professions public may focus attention away from other sources of job satisfaction, and consequently reduce welfare.⁵ As another example, when people make punitive monetary assessments, absent clear reference values, judgments may be diffuse (similar to our findings in Experiment 3), and thus effectively incoherent (cf. Sunstein et al. 2002). Ironically, this can be remedied if reference values (transaction cues) are available and salient, even if those values are only remotely relevant.

Many prominent examples of inconsistency between predicted utility and monetary judgments (such as those mentioned at the outset of this paper) appear in the literature. Our account may help understand such instances. For example, it may help explain why substantial anchoring effects are found with monetary evaluations but not with equivalent pleasure assessments (Ariely et al. 2003). Another class of examples can be traced to undue reference price effects, whereby people are significantly influenced by a salient price that is no longer relevant, such as real estate prices in the city in which they previously lived (Simonsohn and Loewenstein 2006). Similarly, many marketing activities make a reference price salient. Even false messages (cf. Mayzlin 2006) can significantly impact WTP if the consumer is not encouraged to actively consider predicted utility. Our theory speaks to why such effects are observed.

Our account is of particular importance to market researchers who assume that price assessments represent underlying utilities, for example in conjoint analyses; or conversely, assume that eliciting predicted utility reflects WTP. Our results show that this assumption is not always true, and our theory can help predict the likelihood and significance of the disparities. Including price information in a conjoint

⁵ We thank the AE for this idea.

analysis may allow the calculation of optimal price levels, but as our results suggest, it may also lead to overreliance on the transaction and underestimation of the utility associated with other product attributes. Our paper also suggests ways in which the two types of judgments can be streamlined: Before asking for monetary assessments respondents should be encouraged to explicitly consider the utility (e.g., pleasure) they expect to derive from acquiring the purchase alternative. Interventions reminding consumers of the benefits of ownership or consumption (cf. Shiv and Huber 2000) may enhance perceived differentiation and soften competition. Thus, firms may be able to help consumers more fully value products they offer, possibly reducing unfavorable reference price effects.

Our research points to interesting questions for future research. For example, would prior ownership or consumption of evaluated items (cf. Strahilevitz and Loewenstein 1998) help monetary assessments to better reflect predicted utility? Does the manner in which explicit product comparisons influence bidders' reservation prices suggest an increased focus on the predicted experience as opposed to the actual price that would *win* the auction (Dholakia and Simonson 2005)? Another question relates to the impact of availability of cognitive resources. For example, distraction can increase weighting of affective components and decrease that of informational components in assessment of experiences (Nowlis and Shiv 2005). Greater involvement (or greater effort, cf. Kivetz 2003) in monetary evaluation may similarly lead to greater congruency with predicted utility. Another question is whether choices that reflect predicted utility will be more satisfying than choices that do not, as intuition suggests. We suspect that social welfare can be enhanced if consumers base their decisions on experience cues of the type that inform predicted utility, and thus believe that it is important to seek additional ways to reduce the dissociation between monetary assessment and predicted utility.

References

- Ariely, D., G. Loewenstein, D. Prelec. 2003. Coherent arbitrariness: Stable demand curves without stable preferences. *Quart. J. Econom.* **118**(1) 73–105.
- Carmon, Z., D. Ariely. 2000. Focusing on the forgone: Why value can appear so different to buyers and sellers. *J. Consumer Res.* **27**(3) 360–370.
- Dhar, R., S. M. Nowlis. 2004. To buy or not to buy: Response mode effects on consumer choice. *J. Marketing Res.* **41**(4) 423–432.
- Dholakia, U. M., I. Simonson. 2005. The effect of explicit reference points on consumer choice and online bidding behavior. *Marketing Sci.* **24**(3) 206–217.
- Hsee, C. K. 1999. Value seeking and prediction—Decision inconsistency: Why don't people take what they predict they'll like the most? *Psychonomic Bull. Rev.* **6**(4) 555–561.
- Kahneman, D. D., T. Miller. 1986. Norm theory: Comparing reality to its alternatives. *Psych. Rev.* **93**(2) 136–53.
- Kahneman, D., J. Knetsch, R. H. Thaler. 1986. Fairness as a constraint on profit seeking: Entitlements in the market. *Amer. Econom. Rev.* **76** 728–741.
- Khan, U., R. Dhar. 2006. Licensing effect in consumer choice. *J. Marketing Res.* **43**(2) 259–266.
- Kivetz, R. 2003. The effects of effort and intrinsic motivation on risky choice. *Marketing Sci.* **22**(4) 477–502.
- Kreps, D. 1990. *A Course in Microeconomic Theory*. Princeton University Press, Princeton, NJ.
- Mayzlin, D. 2006. Promotional chat on the Internet. *Marketing Sci.* **25**(2) 155–163.
- Nowlis, S. M., B. Shiv. 2005. The influence of consumer distractions on the effectiveness of food sampling programs. *J. Marketing Res.* **42**(2) 157–168.
- Schkade, D. A., D. Kahneman. 1998. Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psych. Sci.* **9** 340–346.
- Shafir, E., P. Diamond, A. Tversky. 1997. Money illusion. *Quart. J. Econom.* **112**(2) 341–374.
- Shiv, B., J. Huber. 2000. The impact of anticipating satisfaction on choice. *J. Consumer Res.* **27**(2) 202–216.
- Shiv, B., Z. Carmon, D. Ariely. 2005. Placebo effects of marketing actions: Consumers may get what they paid for. *J. Marketing Res.* **42**(4) 383–393.
- Simonsohn, U., G. Loewenstein. 2006. Mistake #37: The effect of previously faced prices on current housing demand. *Econom. J.* **116**(1) 175–199.
- Simonson, I., Z. Carmon, S. O'Curry. 1994. Experimental evidence on the negative effect of product features & promotions on brand choice. *Marketing Sci.* **13** 23–40.
- Strahilevitz, M. A., G. Loewenstein. 1998. The effect of ownership history on the valuation of objects. *J. Consumer Res.* 276–289.
- Sunstein, C. R., D. Kahneman, D. Schkade, I. Ritov. 2002. Predictably incoherent judgments. *Stanford Law Rev.* **54**(6) 1153–1215.
- Thaler, R. H. 1985. Mental accounting and consumer choice. *Marketing Sci.* **4** 199–214.
- Trope, Y., N. Liberman. 2003. Temporal construal. *Psych. Rev.* **110** 401–421.
- Tversky, A., S. Sattath, P. Slovic. 1988. Contingent weighting in judgment and choice. *Psych. Rev.* **95**(3) 371–384.