

Decisions by Rules:
Disassociation between Preferences and Willingness to Act

Working Paper
Monday, December 17, 2003

On Amir

Yale University

Dan Ariely

Massachusetts Institute of Technology

On.amir@yale.edu

Yale School of Management

135 Prospect st.

New Haven, CT 06511

The authors would like to thank Orly Lobel, Drazen Prelec, Shane Fredrick, Ely Dahan, Nicholas Epley, Peter McGraw, & Jonathan Levav.

Abstract

Individual decision making has been largely viewed from an outcome maximizing perspective. Building on previous theoretical works, this paper suggests that individuals consider not only their preferences for different alternatives, but also behavioral rules as guiding principals for their choices. It is further proposed that these rules can be invoked by the situation at hand, and therefore may lead individuals to choose alternatives that fit their rules but not their preferences, and thus to not maximize their happiness. A series of experiments demonstrates that individuals follow such rules; that the rules they follow can be disassociated from their preferences, consequently leading to loss of utility; that monetary exchanges can act as invoking cues for such rules; that the use of the rules does not arise because of cognitive limitations; that once the rules are invoked they are used; and that rules are used as a first response to the decision problem and can be overridden. These findings may explain some of the systematic inconsistencies in the ways consumers behave.

Suppose you immensely enjoy smelling the fresh air in the park on the way to work, and even mention this daily to a close friend. Suppose further that a young entrepreneur realizes this, and convinces city hall to start charging one cent for passage through the park, when the flowers are in bloom. How would you react to this new policy? We propose that in such a case, and despite the high utility one would get from the experience (breath of fresh air minus the cost), we will refuse to pay merely because we do not believe one *should* pay for such experiences. This simple example is an attempt to illustrate a simple case where individuals rely on a rule for action rather than on utilities when making choices.

Building on theoretical views in psychology, philosophy, decision making (Prelec 1991; Prelec & Hershstein, 1991; Ainslie, 1992; March, 1994), as well as works by moral and legal philosophers such as Raz (1975), the current work aims to provide evidence for one mechanism that generates preference-action inconsistencies – the use of decision rules¹ – and elaborates on the mechanism underlying the use of rules in making decisions. The mechanism of rules has some resemblance to the use of heuristics, but also some important differences (for a deeper analysis of the differences we refer you to our general discussion). First, heuristics are general procedures used by decision makers to simplify decisions (Frederick, 2002), whereas the rules investigated here provide domain specific “do and don’t do” action plans that are not meant to simplify decisions but rather to enforce certain conventions. Second, while heuristics evolve from tradeoffs between computational demands and accuracy (Simon, 1957; Payne, Bettman, & Johnson, 1993), rules are derived from social, cultural, and general principles, and are thus more similar to moral judgments.

¹ A great deal of discourse in moral and legal philosophy is devoted to the overlap, hierarchy, and boundaries of a myriad of terms describing behavioral guidelines, such as rules, principles, norms, rationales, morals, and so forth. The locus of distinctions between such terms lies in the [superhuman] ability to identify the source, relevance, and distinct implications of each one, contrast, compare, and finally choose the correct manner of thought and action that follow from that choice (See for example, Dworkin’s “Judge Hercules” in Habermas 1996). In this work, the use of the term “rules” refers to a general underlying mechanism whereby people choose according to what they think they ought to do.

The rest of the paper continues with a theoretical background related to the use of rules in decision making, followed by a deeper investigation of one particular rule relating to intertemporal choice – “Don’t pay to delay good experiences.” The mechanism underlying this rule will be investigated through manipulation of the nature and saliency of rule-invoking cues, manipulation of the depth of consideration of the outcome, examining the propensity to override the rule, and by looking at personality differences. Finally, the question of the activation-invocation source of the rule will be addressed. The paper concludes with a general discussion of the specific rule opposing delay of positive experiences as well as the use of decision rules in everyday life.

Why “Rules”?

Following a substantial treatment in philosophy, Prelec (1991; Prelec & Herrnstein, 1991; see also Ainslie, 1992; March, 1994) suggested a decision making style in matters of self-control and identity maintenance that is independent from tastes or preferences but is akin to the use of legal rules. Unlike preferences these rules are assumed to be general over-arching guidelines for behavior and as such are applied broadly in a “do or do not do” legal-like manner – or more generally as stated by Raz (1975) “*x* should [not] do *y* in situation *S*.” Raz (1975) goes further to define a test of whether a behavioral guideline is a rule by identifying its overarching nature and seldom questioned validity:

“[By following a rule] What I am not doing is assessing the merits of the case taking all the relevant facts into consideration. I am not doing this for I have decided on a rule, that is, I have accepted an exclusionary reason to guide my behavior in such cases. I may occasionally, of course, examine the justification of the rule itself. If I re-examine the rule on every occasion to which it applies, however, then it is not a rule which I have adopted. I may on the other hand examine the rule occasionally even when not confronted with a case to which it applies. This is the test by which to determine whether a person follows a rule.”

One consequence of their broad application is that there are circumstances for which the rules are not suited and yet they are applied (as with many legal systems), resulting in actions that

can at times be disassociated from preferences. According to some, it is by the very definition of rules that we exclude a set of considerations that may, at times, prove the right set to apply (Raz, 1975). Recently, researchers have begun uncovering the mechanisms underlying some of these documented inconsistencies. Hsee et al. (2003) describe a set of ‘expressed preferences – anticipated action’ inconsistencies and categorize them into three subgroups, based on the underlying logic (principles) that generates them: lay economism, lay scientism, and lay functionalism. The common thread proposed is that people follow “rationales” – meta-rules or principles that guide behavior and are sometimes applied even when doing so leads to the aforementioned inconsistencies and consequently to less preferred outcomes. The current work follows a similar line of thought focusing on one such rule, and attempts to provide more detailed evidence for the decision making mechanism itself.

Rules in daily decisions

Decision-rules could be based on social, cultural (don’t tell one’s spouse dinner is not good), personal (set a personal deadline one day before the actual deadline), or moral principles (paying for sex is not part of a decent relationship). Growing up in different cultures, or subcultures, individuals are likely to learn different “causal schemata” of the stimulus-response relations in their environment (Kelly, 1972; Nisbet & Wilson, 1977). For example, in his theory of human sociability, Fiske (1992) suggests four meta-schemas of social order that translate to specific courses of action through the adoption of various corresponding norms and rules of conduct by different cultures. The four forms of sociality (communal sharing, authority ranking, equality matching, and market pricing) thus define four distinct sets of admissible behaviors. Such normative behavioral rules could be taught, but could also be learned from experiences (Schwartz, 1977; Gilbert, 1995; see also Raz, 1975), yet vary between individuals within the same society and

function to direct behavior in particular situations (what should one bring as a present for the host of a dinner party). Because these rules are learned and may be socially constructed they are also not universal and the specific rules that are applied in different circumstances depend on individual, social, and cultural factors. Moreover, individuals may “drift” into following rules gradually over a period of time without ever consciously deciding to do so (Raz, 1975):

“A Person may, however, come to follow a rule without having decided to do so. He may have been brought up from early childhood to believe in the validity of the rule and to respect it. He may have drifted into following the rule as an adult gradually over a period of time without ever really making up his mind to do so.”

Several theorists have also argued that personal rules are the predominant mechanism for pre-commitment against follies of self-control by developing general principles, such as “maintain health” or “be good”, effectively uniting actions under a common precept towards a desired end, which later serve them as building blocks of willpower (Prelec, 1991; Prelec & Herenstein, 1991; Ainslie, 1992; Raz, 1975). In these cases, rules may be self-generated in order to dictate future behavior and avoid future conflicts. For example, an ex-smoker may create a rule that prevents her from even an occasional seemingly harmless taste of a friend’s cigarette for fear of increasing her own urge to smoke. The rules under current focus are more implicit in nature, and are thus closer to Raz’s quote than to the writings of Prelec and colleagues (e.g., “I decide upon a rule”).

Rule Activation & Overriding

The activation of such rules is likely not only to be defined by social and cultural context, but also by the local context of the task (March, 1994). In line with this idea, one interpretation of Fiske’s Relational Theory (Fiske, 1992) is that the invocation of money as the central aspect of an exchange will bring about the activation of labor market as a general schema and thus influence willingness to exert effort, reciprocity, and willingness to act selfishly (Gneezy and Rustichini, 2000; Heyman and Ariely, 2003). Similarly, as suggested by Fiske and Tetlock (1997) overriding

rules is also likely to be based on local context, where one could overcome rules emerging from taboo-tradeoffs for example, by obfuscating the tradeoff, and reframing the decision task in more comparable terms (see also Ratner & Miller, 2001, Experiment 4). For example, although it is not considered appropriate to pay with cash for a dinner one is invited to, it is acceptable (and even recommended) to “pay” back the hosts with a gift such as wine or flowers. The key here is that when attention is directed at a rule-invoking cue, the rule will be invoked and followed, and when attention is not directed at a rule-invoking cue, the rule will not be invoked and therefore will not be followed (Ainslie, 1992). Indeed, it has been suggested that the problematic effect of framing on the activation of social norms may lie in the focus of attention (Cialdini, Kallgren, and Reno, 1991). The framing and attentional aspects of the activation of rules also implies that their activation is in many cases implicit, which renders their invocation exogenous, meaning that different market conditions can elicit the use of rules simply based on cue saliency.

Overview

In summary, the rules of current interest are analogues to legal or moral rules - they do not form a real binding constraint on behavior, but instead form guiding principals that are not always explicitly considered. Because these rules form only virtual constraints on decisions, they will be followed most frequently when they are strongly invoked, when decisions are made mindlessly (Langer, 1989), when there is little experience in the domain, or when the roles are acquired via non-personal (social) experience (Reagan & Fazio, 1977). Moreover, because rules are general in nature, behaviors that are guided by them will not always match optimizations according to preferences, and following rules could come at the expense of personal utility.

Within this general framework, the goals of the current work are: to provide further evidence for the reliance on rules as decision mechanisms; to propose and test one particular such rule –

“Don’t pay to delay good experiences;” to demonstrate that reliance on this rules can undermine anticipated utility; to demonstrate that mentioning financial exchanges can by itself invoke reliance on the rule; and to test the mechanism underlying the use of such rules. The strategy used to shed light on the rule as a decision mechanism included several steps. First, one specific rule of consumption was selected– “Don’t pay to delay good experiences.”² Second, in order to be able to test behavior-preference inconsistency in pilot experiments 1 and 2 we looked for a case where preferences are likely to deviate from this rule (i.e., preferences for delay). After finding the rule and establishing a behavior-preference inconsistency (experiments 1 and 2) we move to test 4 aspects of the proposed mechanism: 1) That rule activation is invoked when the decision medium is monetary (Experiment 3). 2) That reliance on the rule is sensitive to the salience of the invocation cue (experiment 2). 3) That careful consideration can negate the exclusionary nature of rule-based decision making (Experiment 4). And 4) that personality differences on a scale related to the tendency to follow rules (self-discipline) predict the preferences-behavior dissociation in face of the rule (Experiment 5).

Pilot experiments 1 & 2

In order to demonstrate that decision makers are not willing to pay for delays, despite the fact they prefer them, one needs to first show that in some instances decision makers have positive preferences toward delays. Likely cases for preference for delay are situations where anticipation can generate positive utility (Loewenstein, 1987). These types of experiences are likely to be long awaited one-time positive events such as a kiss from a movie star or a concert of a favorite band. Given the difficulties in using kisses as an experimental stimuli we picked a concert as the focal

² This rule was selected based on a perceptual map study with 30 MBA students, where the majority of the students viewed (and stated that) the request to pay more for longer delivery periods as unacceptable.

stimuli. The hypothesized pattern of anticipated happiness from such an experience is expected to be significantly different from a product that one is impatient to own and delivers high utility levels over time once owned – such as a new PDA (gizmo). To support this claim, Pilot experiment 1 elicited expected daily happiness for these two products (concert vs. electronic gizmo). The hypothesis was that the temporal patterns of predicted happiness would be different, and in particular that the concert, but not the gizmo will have a large anticipatory component. Pilot experiment 2 examined preference for delay directly by asking respondents for their timing preference for the concert and the gizmo. The hypothesis was that individuals would have a preference to delay the concert, but not delivery of the gizmo. Once the pilot experiments demonstrate a preference for delay for the concert, this stimulus will be utilized throughout the rest of the work to investigate the cases under which this preference for delay will be expressed or suppressed based on the more general rule to avoid paying for delay of positive outcomes.

Pilot experiment 1 Method: Fifty students from a west coast university participated in the experiment as part of a course requirement. Participants were asked to imagine that they have just purchased an item (a ticket to a concert or the latest PDA – a gizmo) and were shown a calendar depicting the next sixty days, with one of three dates marked in bold, one in each condition. The marked dates were tomorrow, two weeks, or one month from today, and represented the actual delivery / consumption time for the given product. This resulted in a six conditions: 2 products (concert vs. gizmo) x 3 delivery / consumption dates (tomorrow, two weeks, or one month) within-respondent design. As the choice was predetermined, participants were asked to report their expected daily happiness (for every day in the calendar, which spanned 2 months) from the

experience described to them. Ratings of daily happiness were recorded on a scale from -10 (unhappy from the experience) to 10 (very happy from the experience).

Pilot experiment 1 Results & Discussion: Figure 1a displays participants' average expected daily happiness. Consistent with the predictions of the positive value of anticipation for the concert, the expected happiness prior to the experience was higher for the concert than for the gizmo, while the expected happiness after the experience was higher for the gizmo than for the concert – the interaction between type of experience and happiness period (before / after) was significant [$F(1,37) = 10.99, p < 0.002$]. In other words, the results show that it is better to have the gizmo earlier, but the same is not true for the concert where as hypothesized there is a substantial preference for delay.

One other way to view these results is to compound the expected daily happiness into a single measure, which presumably is the measure respondents use when making choices between such options. In order to calculate this composite measure we assumed exponential time discounting (daily $\delta = 0.96$), and summed the discounted states across the entire time frame (two months). As can be seen in Figure 1b, the hypothesized translation of expected daily happiness into a single overall satisfaction measure reveals a preference for delaying the concert but not the gizmo (this translation will be tested directly in Pilot Experiment 2)

Overall the results from Pilot experiment 1 show that there exists a class of products (experiences), of which the concert is one example, for which the general rule of not delaying positive outcomes is unsuitable. In fact, if consumers were to use the abovementioned decision rule when making decisions about concert tickets, their actions will not correspond to their anticipated pleasure and result in sub-optimal choices.

••• **Figure 1** •••

Pilot experiment 2 Method: One-hundred students from a west coast university were approached and asked to fill a short questionnaire. Participants were asked to imagine that they have just purchased an item (a ticket to a concert or a gizmo manipulated between respondents) and were asked about their preferences for delivery / consumption time out of four possible dates (tonight, tomorrow, two weeks, or one month ahead). This resulted in a two condition (concert vs. gizmo) between-respondent design. The options for timing (tonight, tomorrow, two weeks, or one month ahead) were displayed on a two month calendar, similar to the one used in Pilot experiment 1, and respondents were asked to indicate the date they desired by circling it on the calendar. In their answers, participants were asked to ignore their schedule constraints. This request is crucial in order to get a clearer measure of participants' preferences for timing because presumably decision makers are more likely to have plans for the near rather than far future.

Pilot experiment 2 Results & Discussion: As seen in Figure 1c, the modal choice for the gizmo was “tonight”, while the modal choice for the concert was “two-weeks from tonight.” Furthermore, the mean desired delay for the gizmo was 1.06 days while the mean desired delay for the concert was 7.25 days [$t(98) = 6.07, p < 0.0001$]. It is interesting to note that for the concert there were two respondents who preferred to delay the experience by a whole month. Examining these results, it is reassuring to note the similarity between Figure 1b (the composite measure based on anticipation), and Figure 1c (choice proportions), which were high for both the gizmo ($r = .99$) and for the concert ($r = .89$). Furthermore, these high correlations held for any daily time discounting level below 0.98 for the gizmo (low discount factors correspond to high time discounting, signifying impatience), and for any daily discount factors greater than 0.94 for the concert (signifying patience).

In summary, the results from Pilot experiment 2 strengthen the conclusion from Pilot

experiment 1 whereby for certain classes of experiences (such as concerts) consumers prefer delay. Given these results the experiments to be reported next all use the concert as a stimulus, testing the conditions under which respondents will reveal a positive preference for delay (consistent with their preferences) and the conditions under which respondents will reveal a negative preference for delay (consistent with the use of the general rule “Don’t pay to delay good experiences”).

Experiment 1 – Preference-WTP Dissociation

Pilot experiments 1 and 2 demonstrate that people would rather attend a concert at a later date (as long as it is not too late) and that this preference corresponds to the happiness they expect to derive from savoring such experiences (Loewenstein, 1987). The next question to ask is whether the contradiction between preference for delay in this specific case (the concert) and the more general rule against paying for delay will cause decision makers to behave in a way that is not consistent with their preferences. The main hypothesis is that when actions such as willingness to pay are concerned, this rule will tilt the scale and decision makers will act (pay) according to the rule and not according to their preferences.

Method: Two hundred eighty-nine students from a west coast university were presented with the same concert descriptions (but not the gizmo) and calendar as in the pilot experiments. The design for this experiment was a 2 (timing) x 2 (type of DV) x 2 (explicitness). The timing factor was manipulated by marking the time of the concert as “tonight” or as “in two weeks.” For the DV factor half of the participants were asked for their willingness to pay (WTP), and half were asked for their predicted total expected satisfaction from the concert on a visual analog marked at one extreme as “extremely unhappy,” at the other extreme as “extremely happy,” and at the middle as “indifferent.” The visual scale was later translated to a 21 point scale ranging from -10 to +10. Note there were no negative responses on this scale, and thus the meaningful scale is bounded between 0

and 10. The explicitness factor manipulated whether respondents were asked to provide expected daily happiness for each of the days in the calendar (as in Pilot experiment 1) or not.

The hypotheses were that while the expected satisfaction will show a preference for delay, the WTP measure will not, and that this tendency will become stronger when providing explicit daily happiness as this manipulation will make the anticipated pleasure clearer for the participants in the expected satisfaction but will not change the response for those following the rule (WTP).

Results & Discussion

The satisfaction results generally replicated the pilot experiments. Overall, expected satisfaction was higher for the delayed date (mean = 7.05) than for the near one (mean = 5.43), and this difference was statistically significant [$F(1,136) = 16.05, p < 0.001$]. The WTP results were in the opposite direction, where overall, WTP was lower for the delayed date (mean = 25.17) than for the near one (mean = 30.42), and this difference was statistically significant [$F(1,144) = 7.01, p = 0.009$].

Aside from these two main effects, it is interesting to note how the request to provide expected daily happiness measures influenced respondents' estimates. As can be seen in Figure 4, the total expected satisfaction was clearly influenced by the request to provide daily-expected happiness ratings. In this case, total happiness increased when respondents became more aware that in this special case, there is a benefit for delay, leading to a significant interaction [$F(1,136) = 4.69, p = 0.032$]. On the other hand, the WTP measures were not influenced by the request to provide daily-expected happiness ratings [$F(1,144) = 0.19, ns.$]. The results are consistent with the idea that when asked to provide a payment (or a WTP response) the rule against payment for delay is still invoked so that even making the benefits from delay more apparent did not change the pattern of the results. Together, the lack of influence of providing expected daily happiness on WTP and its

influence on total happiness, create an increased disassociation between preferences and WTP as the awareness of the benefits of delay increase.

••• **Figure 2** •••

In sum, the results of Experiment 1 show that decision makers are not always able to correctly relate their actions (WTP) to their preferences (happiness). Moreover, when reminded of their preferences, individuals display an even greater discrepancy between expected satisfaction with delay and willingness to pay for it – suggesting that increasing individuals’ understanding of their preferences does not always free them to act in their best interests. This increased discrepancy because of the performance of a simple prediction task indicates that the preference - WTP discrepancy is not due of a heuristic aiming to minimize cognitive effort, but instead it may be based on a strong behavioral rule —not to pay for delay. If such a heuristic existed and was the cause of the discrepancy, the prediction task should have decreased the discrepancy.

Experiment 2 – Invocation Saliency

The WTP measure used in Experiment 1 has some drawbacks associated with it. First, WTP responses are bounded from below and unbounded from above. Second, there are concerns about participants’ ability to provide meaningful responses to this type of WTP question (Fischhoff, 1991). To address these concerns, an attempt was made to replicate these results with a choice measure, which is not susceptible to scale interpretations. More important, Experiment 2 also tested more directly the idea of rule-based choice by manipulating the saliency of the rule – hypothesizing that as the rule becomes more salient, it will be followed more frequently, leading to a larger

disassociation between choices and preferences.

Method: The questionnaire used had the same basic concert scenario as in Experiment 1. Two hundred and forty-six students from a west coast university were approached and asked to fill a short questionnaire. The respondents were randomly assigned to one of four DV conditions in which they were asked to chose their preferred timing for the concert ('tonight', 'two weeks from tonight,' or indifferent) such that this choice will maximize their: pleasure, preference, WTP, or paying for switching (which would you pay more for: to change your ticket from tonight to two weeks or from two weeks to tonight).

It was hypothesized that the rule prohibiting payment for delay (and thus creating what seem to be a preference for immediacy) was least salient in the pleasure condition, more salient in the preference condition, even more salient in the WTP condition, and most salient in the most explicit case –switching. Accordingly, preference for delay would seem strongest in the pleasure condition, followed by the preference condition, followed by the WTP condition, with the switching condition showing the least preference for delay.

Results and Discussion

The results confirmed the hypothesis (see Figure 3). Preference for delay was strongest when respondents made choices according to their pleasure (58%), weaker in preference (48%), and much weaker for WTP (19%) and switching (21%) where there was a substantial preference for immediacy. The three-level choice data was further analyzed using an ordered probit model³, in which each condition was represented by a dummy variable. In this analysis, a negative coefficient represents preference for immediacy (impatience), a coefficient not different from zero represents indifference, and a positive coefficient represents a preference for the delayed concert (patience).

³ In all following analyses, reducing the data to a binary measure by either splitting indifference into the two polar choices or by ignoring it yields qualitatively similar results.

The ordered probit analysis confirms that relative to the preference condition, participants were more likely to choose the delayed concert in the pleasure condition [$\beta = 0.355$, $t(382) = 1.86$, $p = 0.063$], but more importantly, participants in the payment condition and in the switching condition were all much more likely to choose the immediate concert [$\beta = -0.446$, $t(382) = 2.38$, $p = 0.017$, and $\beta = -0.531$, $t(382) = 2.08$, $p = 0.038$, respectively].

In order to further understand the underlying time preference, for each of the conditions the number of participants who chose to see the concert in two weeks was divided by the number of participants who chose to see the concert the same night. Using this ratio measure, a score that is greater than 1 indicates preference for delay, smaller than 1 indicates preference for immediacy, and 1 indicates indifference. The results confirmed our expectations where all ratios but the one for the preference question were statistically different from 1 ($p < 0.001$). When respondents were thinking about their pleasure, there was a strong preference for delay (2.07), which was reversed when asked about WTP (0.36) or paying explicitly for switching (0.27), while the preference question showed an even preference between the two options (1.04). The preference results could suggest that when asked about their preferences in a generic framing, some decision makers use WTP type of reasoning and others use pleasure type of reasoning when answering. Overall, these results support the “decisions by rules” framework in two ways: First, the results replicate the disassociation between pleasure and payment, and second they show that as the rule becomes more explicit, such as in the switching condition, the disassociation between preferences and action becomes even larger.

••• **Figure 3** •••

Experiment 3 – Invocation Cue

The disassociation between stated preferences and stated WTP observed in the preceding experiments is argued to stem from the [mis]use of rules. In particular, it was hypothesized that when it comes to payment, individuals follow a rule that at times leads them towards an inferior outcome. This discrepancy between participants' behavior and their best interests may be specific to money (as is currently hypothesized) or may be an outcome of taking an action (such as payment) or a bias in the process of predicting happiness or future actions. One way to answer this question is to compare WTP responses to other proxies for preference that involve action and future predictions but not money. The proxy that is closest to money without being money seems to be willingness to exert effort (WTE). The current hypothesis is that thinking about money is the cue that invokes the rule. If this is correct, WTE should give rise to a similar pattern of results as the pleasure questions but different from WTP. Conversely, if WTE will give rise to results similar to WTP and dissimilar from pleasure, we would strengthen our belief that the cause of the disassociation lies in other more general response mode effects.

Method: Experiment 3 replicated the basic question of the concert timing choice (tonight, two weeks from tonight, or indifference), using three different dependent measures (WTP, happiness, and WTE) in a between-participants design. In the WTP condition participants were asked which timing option they would be willing to pay more for; in the happiness condition participants were asked which timing option would give them greater overall happiness; and in the WTE condition participants were asked for which timing option they would be willing to drive farther in order to purchase the tickets today (driving to buy tickets today was important in order to control for the temporal effect on the perceived effort involved). One hundred and eighty-four students agreed to answer a short questionnaire.

Results & Discussion

As can be seen in Figure 4a, the results support the idea that rules are activated when monetary related judgments are called for, sometimes overriding preferences. The same type of probit analysis used earlier showed that asking about WTP as opposed to happiness replicated previous findings, whereby participants indicated that they will enjoy the delayed concert more but would pay more for the earlier one [$\beta = -0.495$, $t(181) = 2.38$, $p = 0.017$]. More central to the current experiment, when asked about the concert for which they will drive more to get the ticket for, participants were as patient as in the happiness condition [$\chi^2(1) = 0.45$, ns.], and significantly more patient than the WTP condition [$\beta = 0.633$, $t(181) = 2.98$, $p = 0.003$].

In sum, decisions regarding willingness to exert effort were consistent with pleasure judgments but inconsistent with willingness to pay judgments. These results suggest that the WTP-predicted preference inconsistency demonstrated here is not due to the action implied by monetary decisions or to a bias in the process of predicting happiness or future actions. Instead, it seems that the rule for action used here is invoked when individuals think about money.

••• Figure 4 •••

Experiment 4a – Consideration Override

The results of the previous experiments support the proposed mechanism by demonstrating that in some cases individuals make choices that are not in line with their preferences and that this tendency is invoked by making money-related judgments. We take this as initial evidence that when it comes to payments individuals sometimes follow rules for action in a moral-like manner

(i.e. do the right thing as opposed to generate the most “good”). One other aspect of rules as a decision mechanism is that individuals may mindlessly follow rules, as suggested by Raz (1975) (See also Langer 1989). If this is indeed the case, then asking people to consider the decision more carefully should alleviate the inconsistency between predicted happiness and WTP (a “system 2” override, Kahneman & Fredrick, 2003). A second approach for overriding such rules is one that makes individuals more cognizant of their hedonic needs. Experiments 4a and 4b tested these two override conjectures respectively.

Method: Three hundred and thirty seven individuals participated in a short online experiment, replicating the basic question of the concert timing choice (tonight, two weeks from tonight, or indifference), with one of two different dependent measures (WTP and happiness). The new manipulation in this experiment was “reaction time,” where half the participants were asked to answer with the first response that comes to mind, and the other half was asked to answer only after carefully considering the question. This resulted in a 2 (WTP vs. Pleasure) x 2 (first reaction vs. careful) between-participants design with concert timing (tonight, for two weeks from tonight, and indifference) as the main dependent measure and time to make the decision as a manipulation check.

Results & Discussion

In term of manipulation check, the amount of time respondents took to answer the question was higher when instructed to answer carefully compared with the time it took them to answer when instructed to use their first impression [$F(1,336) = 6.26, p = 0.0128$].

We next examined whether the results replicated previous findings (see Figure 4b) by using the same ordered probit analysis which included dummies for the effect of the different dependent measures levels. The happiness WTP disassociation replicated previous results [$\beta = -0.224, t[336] =$

1.82, $p = 0.069$] predicting greater pleasure in delayed concerts but willing to pay more for the earlier ones.

Finally, we turn to the gist of the current experiment by testing whether the thoughtfulness manipulation influenced the propensity to use the rule. In order to do so we estimated the interaction between consideration level and DV (WTP vs. Pleasure) which was significant and negative [$\beta = -0.333$, $t[335] = 2.21$, $p = 0.027$]. This interaction points to a higher likelihood to follow the rule when indicating immediate reaction compared with more careful judgments – meaning that the first reaction is to follow rule and preferences can override it later if more thoughts are directed to the task.

The overall picture from Experiment 4a is that once a rule is invoked individuals follow it without engaging in deep considerations of its appropriateness (Raz 1975; Langer 1989) unless they pay more attention to the decision problem at hand. It is also worth pointing out that additional thinking did not influence the reaction to the happiness question, presumably because this is an easy judgment and more thinking does not influence it, but this additional thinking did change the judgments regarding WTP, presumably because it facilitated overriding of the initial rule.

Experiment 4b – Hedonic Override

Another situation which may encourage overriding rules is one in which the hedonic (pleasure) considerations of the decisions are more salient. Experiment 4b tested this idea by priming the hedonic aspects of the decision.

Method: One hundred and forty-five undergraduates at an east-coast university participated in the experiment in return for candy. The central manipulation involved priming half of the participants with hedonic thoughts by asking them to write about their experiences when eating

their favorite foods or when having a full-body massage before making their choice in the main task. Thus, the design was a 2 (hedonic priming vs. no priming) x 2 (WTP vs. happiness) in a between-participants design. The main dependent measure was the timing choice for the concert.

Results & Discussion

As in previous experiments, participants' choices were subjected to an ordered probit analysis, which included dummies for the effect of the WTP measure and for the priming manipulation (choice proportions are displayed in Figure 4c). The main effect of the dependent measure replicated previous findings by showing a significant negative coefficient [$\beta = -0.706$, $t[141] = 4.21$, $p < 0.001$]. This means that when asked for the concert date for which participants would be willing to pay more, as opposed to the date in which they would enjoy the concert more, participants were much more likely to choose the earlier date than the later one. Examining the effects of the priming manipulation showed that hedonic priming had a significant positive effect on participants' choice of dates (induced patience), relative to the control where there was no priming manipulation [$\chi^2(1) = 6.70$, $p = 0.009$].

In sum the results from Experiment 4b strengthen the results of Experiment 4a by providing converging evidence for rule override. In particular Experiment 4b demonstrates that making participants more aware of their hedonic needs decreased their tendency to follow the rule.

Experiment 5 – Individual Differences

The results of the previous experiments support the proposed mechanism on an aggregate level. Moving to an individual level analysis would increase our understanding of the specific phenomenon, but more importantly, may provide additional support for the rule-following mechanism. The approach we adopt here is to examine the relationship between the preference-

WTP discrepancy in concert timing choice and one of the “big-five” (Costa & McCrae, 1988) personality traits: conscientiousness, which includes the tendency to obey and follow rules (self discipline). The hypothesis is that higher reliance on the rule (as exhibited by preferences of higher WTP for earlier concerts) will be associated with higher scores on the conscientiousness scale.

Method, Results & Discussion

One hundred and nineteen participants were asked for their concert timing choice (tonight, two weeks from tonight, or indifference), in one of the two basic dependent measures (WTP and happiness). After providing their response, participants were asked to answer the TIPT (Gosling et al. 2003), which is a 10-item, 7-point, measurement scale of the Big-Five personality dimensions.

Similar to previous experiments, participants’ choices were subjected to an ordered probit analysis, which included a dummy variable for the condition, as well as the personality scale and their interaction. The main effect of asking about WTP as opposed to happiness replicated previous findings [$\beta = -0.626$, $t[118] = 3.00$, $p = 0.003$] demonstrating that participants predicted they will enjoy more the later concert but they were willing to pay more for the earlier one.

Most important for the current experiment, the role of personality traits confirmed the hypothesis (see Figure 5). In the WTP condition participants who chose the immediate concert scored higher on conscientiousness scale compared to the participants who chose the delayed concert [$\beta = -0.412$, $t[107] = 2.13$, $p = 0.033$]. In the pleasure condition participants who chose the delayed concert scored higher on conscientiousness scale compared to the participants who chose the immediate concert [$\beta = 0.270$, $t[107] = 2.12$, $p = 0.034$].

These results support the idea that the choice of immediate and delayed concert, and in particular the preference reversal that was observed across all the experiments presented earlier is connected with the tendency to obey and follow rules. As predicted, while in the pleasure condition

a higher tendency to obey and follow rules was associated with a choice to delay the concert, in the payment condition the opposite was true and a higher tendency to obey and follow rules was associated with a choice to prefer the earlier concert. More generally, the hypothesis for Experiment 5 postulated that personality differences in conscientiousness will be linked to the tendency to follow the “Don’t pay to delay good experiences” rule and the results (see Figure 5) support this hypothesis.

••• Figure 5 •••

General Discussion

The objective of this work was to demonstrate that individuals sometimes make decisions according to preset rules and not their preferences, and that such a decision-making mechanism may lead them to make decisions that are not in their best interests. The general perspective on rules is that they are overarching guidelines for behavior that are learned either from experience or from social exchange and are followed when invoked. A set of experiments focused on one such rule (“Don’t pay to delay good experiences”). These experiments provided support for the decision by rules mechanism by showing that there are cases in which people choose outcomes that are not the outcomes that they prefer, that this discrepancy is caused by following a rule, and that this rule is invoked by monetary considerations.

Pilot experiments 1 and 2 demonstrated that there are at least some desirable experiences (concerts being one of them) for which delayed consumption increases pleasure. Using a concert as a stimulus, Experiment 1 demonstrated a payment-pleasure disassociation where the delayed concert was expected to bring about more pleasure but respondents were willing to pay less for it.

Experiment 2 provided further support for the proposed mechanism by demonstrating that the magnitude of this disassociation was related to the salience of the rule-invoking cue. Experiment 3 provided further evidence for the importance of the rule-invoking cue using a different method of expressing preferences, exerting effort. When asked about effort, participants did not exhibit the payment (in effort)-pleasure inconsistency, but in fact were consistent with the preference for the delayed concert. This result further supports the contention that money is the invoking cue for the rule prohibiting payment for delay; and that, absent that cue, participants remained true to their preferences. Experiments 4a and 4b provided some evidence regarding the thought process that is involved. Experiment 4a demonstrated that the default is to apply the rule intuitively unless asked to carefully consider the decision, at which point the rule is likely to be overridden. Similarly, Experiment 4b demonstrated that when participants were primed to think more about their hedonic states they were also more likely to override the rule. Finally, Experiment 5 provided additional support for the underlying mechanism by demonstrating that individuals with greater tendency to follow rules have a greater propensity to exhibit the abovementioned disassociation.

In summary, the common view held by both individuals and decisions scientists is that individuals make decisions according to a set of preferences by searching an optimum, a local optimum, or for a close enough estimate when exact algorithms are too costly (use of heuristics). In contrast, the current work suggests that this is not always the case and that decision makers sometimes do not try to consider the best alternative according to their preferences, but rather act upon pre-imposed decision-rules that are based on moral or social norms and on behavioral guidelines (see also Prelec, 1991; Prelec & HERNSTEIN, 1991; Ainslie, 1992; March, 1994; Hsee et al, 2003).

Related perspectives

The “decisions by rules” perspective presented here is related to other theories of individual decision making: reason-based choice and heuristics. In this section, we elaborate on these two theories.

According to the reason-based choice view (Shafir et al. 1993), people rely on reasons to justify choices, especially in the face of conflict, and sometimes even search for them when the reasons are not obvious – “people search for a compelling rationale for choosing one alternative over another.” While reasons and rules can be considered close relatives, there are some important differences. Perhaps the most prominent is the hierarchical asymmetry: rules can provide and function as reasons, but reasons seldom become rules (but see Raz, 1975 for an in depth discussion on a hierarchy of reasons, some of which may be used as the rules defined in the current context). Another difference lies in the hypothesized process: The current investigation demonstrates that rules are used on the fly at a low level of thoughtfulness while the analysis of reason-based choice envisions an elaborate process of weighting and comparing reasons pro and con for each available choice alternative. Finally, the situations investigated in the current work do not lend themselves as naturally to a reason-based process, but nevertheless the use of rules would certainly provide easy ex-post reasons to explain the choices participants made.

From the heuristics perspective, there are clearly some similarities to the “decisions by rules” mechanism, but also some important differences. The first main difference concerns the goal of the heuristics / rules. Heuristics are procedures used by decision makers to limit the amount of information processed or the complexity of the ways in which it is combined (Frederick, 2002). As such, heuristics are useful for simplifying computations under uncertainty, when cognitive resources are scarce, or when full computation is infeasible. Heuristics such as elimination by aspects

(Tversky, 1972), and representativeness (Kahneman & Tversky, 1982) are examples of such mechanisms. The rules investigated here are different because they do not describe a computational approach. Instead, these rules provide “do and don’t do” action plans – that are not meant to simplify decisions but rather to enforce certain conventions. The second main difference is related to preferences. Heuristics are said to “work” at the service of preferences – aimed at maximizing them under a certain set of constraints (cost of thinking, time, effort). There are multiple examples of such heuristics based tradeoffs, perhaps the most notable is the accuracy-effort tradeoff framework (Payne, Bettman, & Johnson, 1993). In contrast rules are not related to a sacrifice of utility for the purpose of local effort-benefits considerations. For example, consider a choice between an option that is a “better deal” (e.g. two products cost the same, but one had a higher original price) and one that provides higher personal utility (e.g. the product that is liked more). Under such conditions Hsee (1999) has demonstrated that individuals are likely to chose the “better deal” option implying that they follow the “value seeking” rule rather than their own preferences (for a detailed treatment of this rule, see Hsee, 1999; also Hsee et al., 2003). Thus, rules will be used for guiding decision making even when the complexity of the task is trivial, and when there are no repercussions from not doing so.

Other Decision Rules

The general perspective of the current work is that many decisions are carried out based on pre-existing rules as general mechanisms to guide decisions in day-to-day life. A few examples of such rule-based behavior might be found in domains such as: 1) moral - taking home company office supplies is ok, while taking money from the petty cash box is prohibited; 2) volunteering - people are more likely to volunteer outside their professional domain than within it; 3) purchasing - coupons are acceptable price discrimination but online price discrimination is not (such as in the

Amazon case Rosencrance, 2000); 4) negotiations - people offer and expect more than the equilibrium divisions in ultimatum games (see Roth, 1995 for a review; Frederick, 2002); 5) etiquette - holding the fork in one's left hand is correct even if it is not comfortable; 6) valuation – making choices according to perceived monetary value as opposed to true preferences (Hsee, 1999; Hsee et al., 2003); 7) fairness – it is unfair for coke to charge more for a cold drink on a warm day (King & Narayandas, 2000); 8) inter personal relations - people find it objectionable to pay others to smile or be nice to them, and also prefer paying a babysitter not in front of their children.

This latter rule can be generalized into the idea that it is considered immoral to pay directly for relationships (romantic relationship, buying friends, or hiring people to give one compliments). To demonstrate this rule within the “decisions by rules” framework, one additional experiment was carried out. In this experiment respondents were asked to assume they were going to purchase a TV with a budget of \$500 and were asked to choose their preferred store out of two possibilities. There were three between respondent conditions, which varied on the saliency of paying for relationships. The hypothesis was that as the saliency of paying for relationship will increase the attractiveness of that option will decrease. In the low-saliency condition respondents were given a choice between store A that had average price and average service, and store B that had better service but prices that were 10% higher. In the mid-saliency condition the expensive store (B') had better service but charged 10% commission for their service. In the high-saliency condition the expensive store (B'') had better service but charged \$50 for service, paid at checkout to the salesperson. Note that in all three conditions, the store with the high quality service (B, B', and B'') entailed an identical premium but different framing of the payment for service.

One hundred and forty-six students from a west coast university were randomly assigned to one of three conditions. As hypothesized, the proportion of participants who chose the expensive-

high quality-of-service store was 57% in the low-salience condition, and 28% in the high-salience condition [$t(93) = 3.03, p < 0.004$]. The proportion of participants who choose the expensive store in the mid-salience condition was between the other two conditions (41%). These results showed that although individuals indicated they were willing to pay implicitly for service (prefer a store that is more expensive only because it offers better service), their willingness to pick that store largely diminished as the transaction was framed more explicitly as payment for service. These results provide additional support to the idea that the more salient the rule invoking cues are (as the payment gets more implicit), the more likely are decision makers to follow the rule. Finally, this experiment also illustrates that masking the conflict caused by the decision rule (by bundling for instance) yields choices that are more aligned with preference and are thus normatively better.

Future Research

The current research leaves many open questions for future study: What is the origin of rules such as the one against paying for delay. It is possible that this rule stems from the need for immediate gratification? From a developmental perspective the human ability to anticipate develops at a later stage than the need for immediate gratification and that an extensive literature on delay of gratification finds that the ability to delay gratification is positively related both to developmental level and to surrounding social habits (Mischel, et al, 1989). What other rules are there? What is the process by which these rules emerge? What are the individual and cultural determinants of such rules (see also Fiske, 1992; Fiske and Tetlock, 1997)? Are these rules specific to exchanges with money or do they extend to non-monetary exchanges? What characterizes the set of rules that are invoked by money? Finally, it is interesting to ask whether there are ways in which individuals can overcome the rules and behave according to their preferences. Regarding this last question, it is possible that being mindful (see Langer, 1989) of the different rules might allow consumers to

ignore them at the time of action or to pre-commit to ignoring them ex-ante. Some theorists believe that violating a rule can be a slippery-slope to their demise (see Ainslie, 1992, for a review) while others believe that such violations can be highly informative to the decision maker about her own persona (Prelec, 2003). It may be the case that some types of rules present the former case and others the latter. Resolving this dilemma is at the heart of understanding the use of rules in guiding behavior.

References

- Ainslie, George (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person*. Cambridge University Press.
- Cialdini, Robert B., Kallgren, Carl A. and Reno, Raymond R. (1991). A focus theory of normative conduct: a theoretical refinement and reevaluation of the role of norms in human behavior. In Zanna, Mark P. (Ed.), *Advances in experimental social psychology*, vol. 24, 201-234.
- Costa, Paul T., & McCrae, Robert R. (1988). From Catalog to Classification: Murray's Needs and the Five-Factor Model. *Journal of Personality and Social Psychology*, Vol. 55, No. 2, 258-265.
- Fischhoff, Baruch (1991). Value Elicitation: Is There Anything in There? *American Psychologist*. 46, 835-847.
- Fiske, Alan Page (1992). The Four Elementary Forms of Sociality: Framework for a unified theory of social relations. *Psychological Review*, 99, 689-723.
- Fiske, Alan Page & Tetlock, Philip E. (1997). Taboo Trade-offs: Reactions to Transactions that Transgress the Spheres of Justice. *Political Psychology*, 18, 255-297.
- Frederick, Shane (2002). Automated Heuristics. In Gilovich, Thomas, Griffing, Dale, and Kahneman, Daniel (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press.
- Gilbert, Daniel T. (1995). Attribution and Interpersonal Perception. In A. Tesser (Ed.), *Constructing social psychology* (pp. 98-147). New York: McGraw-Hill.
- Gneezy, U. and A. Rustichini (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics* 115(3): 791-810.
- Gosling, Samuel D., Rentfrow, Peter J., & Swann, William B. (2003). A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality*. In Press.
- Habermas, Jurgen (1999). Dworkin's Theory of Law. In Jurgen Habermas, *Between Facts and Norms*, MIT Press, Cambridge, MA.
- Heyman, James E., & Ariely Dan. Effort for Payment: A Tale of Two Markets. Manuscript Massachusetts Institute of Technology
- Hsee, Christopher K. (1999). Value Seeking and Prediction – Decision Inconsistency: Why Don't People Take What They Predict They'll Like the Most? *Psychonomic Bulletin & Review*, 6 (4), 555-561.

- Hsee, Christopher K., Zhang, Jiao, Yu, Fang, & Xi, Yiheng (2003). Lay Rationalism and Inconsistency between predicted Experience and Decision. *Journal of Behavioral Decision Making*, 16, 257-272.
- Kahneman, Daniel, & Frederick, Shane (2003). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In Gilovich, Thomas, Griffing, Dale, and Kahneman, Daniel (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press.
- Kahneman, Daniel, & Tversky, Amos (1982). Subjective Probability: A Judgment of Representativeness. In Daniel Kahneman, Paul Slovic, & Amos Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- King, Charls, & Narayandas, Das (2000). Coca-Cola's New Vending Machine (A): Pricing To Capture Value, or Not? *Harvard Business School Case # 9-500-068*.
- Langer, Ellen. J. (1989). Minding Matters: The Consequences of Mindlessness-Mindfulness. In Leonard Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 22, pp 137-174). San Diego, CA: Academic Press.
- Loewenstein, George (1987). Anticipation and the Valuation of Delayed Consumption. *The Economic Journal*, 97: 666-684.
- March, James G. (1994). Rule following. In *A primer on decision making: how decisions happen*, The Free Press (57-102), NY, NY.
- Mischel, Walter, Shoda, Yuichi, and Rodriguez, Monica L. (1989). Delay of Gratification in Children. *Science*, Vol. 244, 4097, 933-938.
- Nisbett, Richard E., & Wilson, Timothy D. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, 84, 231-259.
- Payne, John W., Bettman, James R., and Johnson, Eric J. (1993). *The Adaptive Decision Maker*. New York: Cambridge University Press.
- Prelec, Drazen (1991). Values and Principales: Some Limitations on Traditional Economic Analysis. In Etzioni, A. and Lawrence, P.R. (Eds.), *Socio-Economics toward a new synthesis*, M.E. Sharpe, Inc, 131-145.
- Prelec, Drazen (2003). Rules (are ment to be broken). MIT Working Paper.
- Ratner, Rebecca K. and Miller, Dale T. (2001). The Norm of Self-Interest and its Effects on Social Action. *Journal of Personality and Social Psychology*, Vol. 81, 1, 5-16.

- Raz, Joseph (1975). Reasons for Action, Decisions and Norms. *Mind*, Vol. 84, 336 (October), 481-499.
- Reagan, Dennis T. and Fazio, Russel (1977). On the Consistency Between Attitudes and Behavior: Look to the Method of Attitude Formation. *Journal of Experimental Social Psychology*, 13, 28-45.
- Rosencrance, Linda (2000). Customers Balk at Variable DVD Pricing on Amazon.com, *CNN.com.technology>computing*, Online article, Sep. 13th.
- Roth, Alvin E. (1995). Bargaining Experiments. In Kagel, J.H. and Roth, A.E. (Eds) *Handbook of experimental economics*. Princeton university press, NJ.
- Shafir, Eldar, Simonson, Itamar, & Tversky Amos (1993), Reason-Based Choice. *Cognition*, Vol. 49, 11-36.
- Shugan, Steven M. (1980), "The Cost of Thinking," *Journal of Consumer Research*, Vol. 7, No. 2, 99-111.
- Simon, Herbert A. (1957). Rational Choice and the Structure of the Environment, *Models of Man*, Wiley, New York.
- Tversky, Amos (1972), Elimination by Aspects: A Theory of Choice. *Psychological Review*, 79, 281-299.

Figure captions:

Figure 1: a) Average expected daily happiness from the concert and gizmo. b) Translation (sum of exponentially discounted daily values) of expected happiness (panel a) to overall preference. c) Choice proportions among times for the concert and gizmo.

Figure 2: Willingness to pay (a) and expected satisfaction (b) for the two time frames, with and without explicit reporting of day-to-day happiness.

Figure 3: Proportion of choices for the two timing options and the four dependent measures.

Figure 4: Timing preference as a function of a) the different dependent measures in Experiment 3, b) the different dependent measures and level of thoughtfulness in Experiment 4a, c) the different dependent measures and hedonic priming in Experiment 4b. For simplicity this figure present the two timing choices without the indifference.

Figure 5: Average personality ratings, across the two choices and the two conditions.

Figure 1a

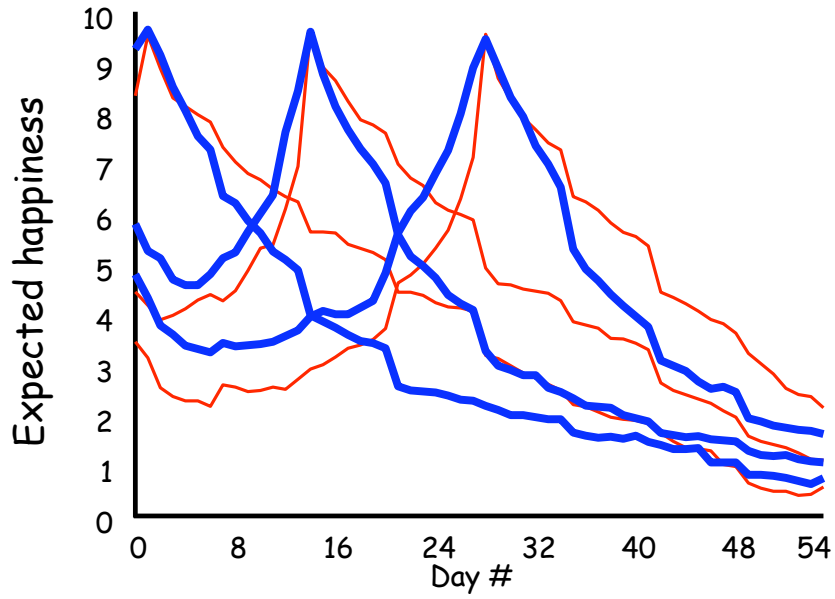


Figure 1b

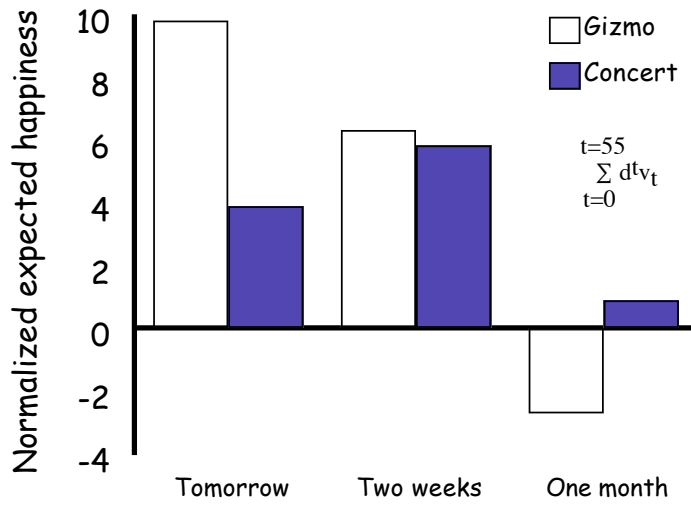


Figure 1c

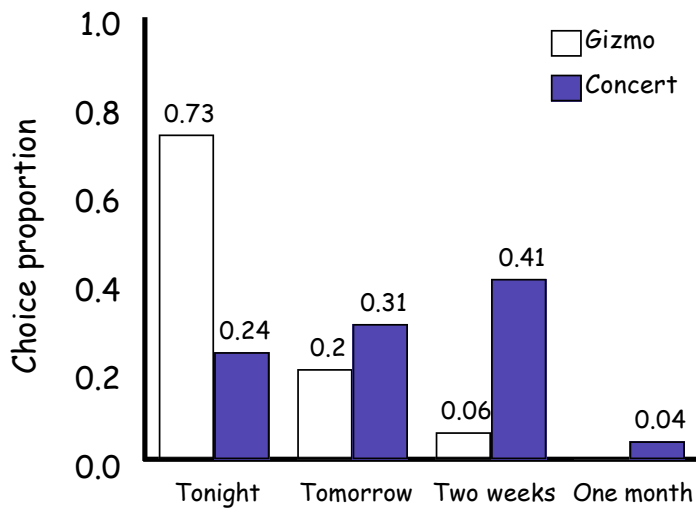


Figure 2a

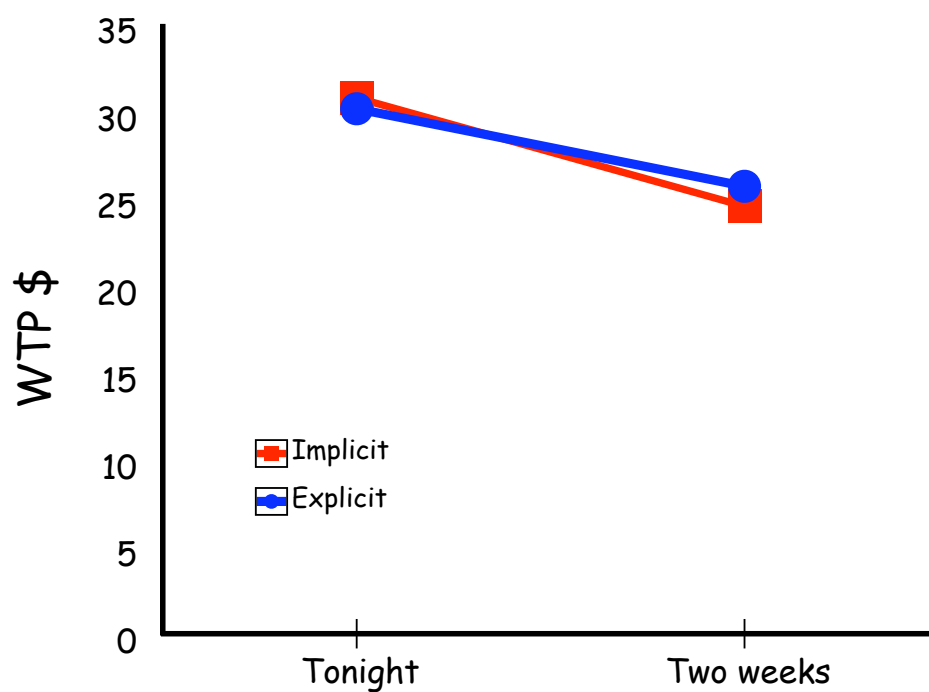


Figure 2b

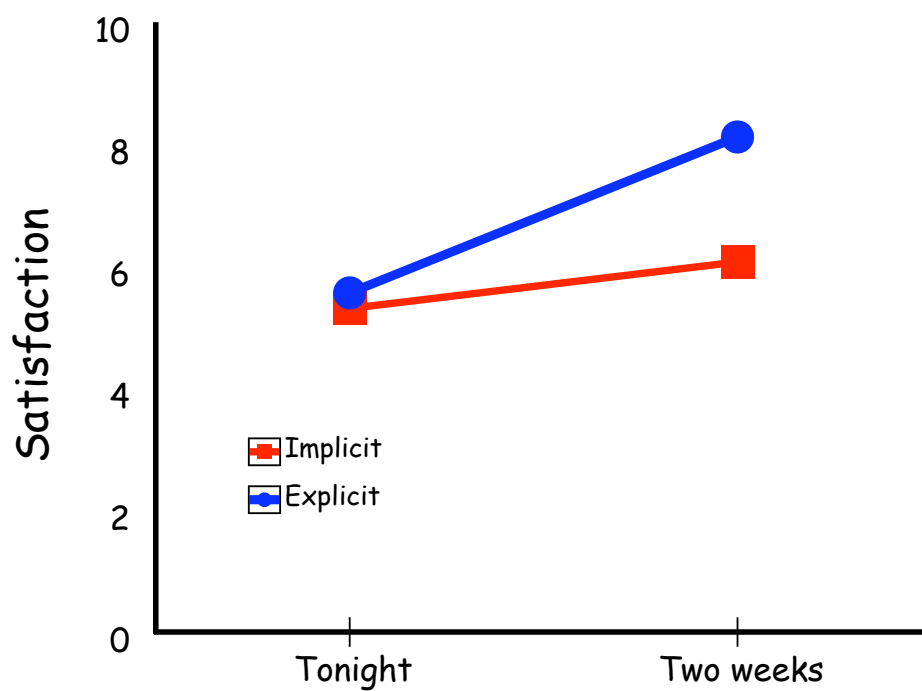


Figure 3

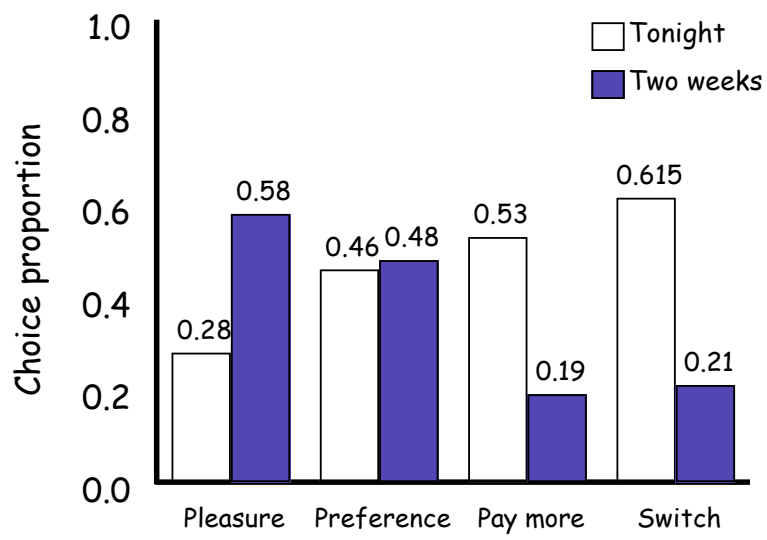


Figure 4a

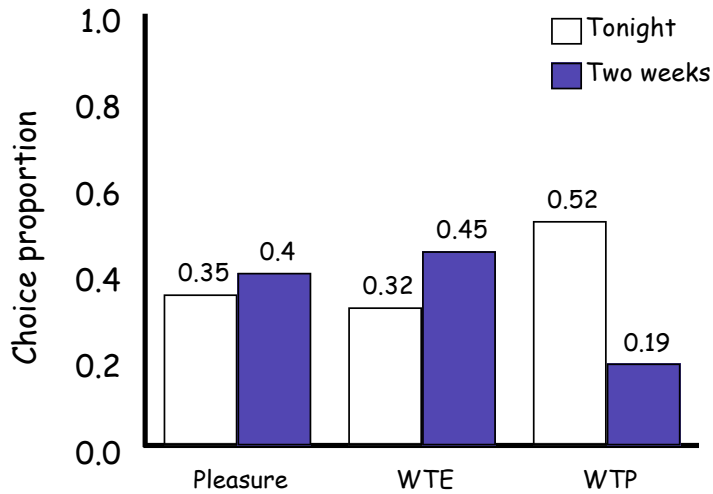


Figure 4b

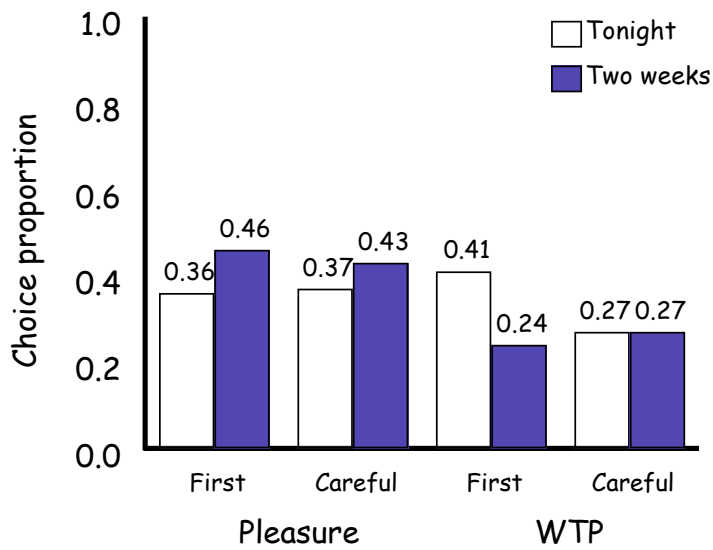


Figure 4c

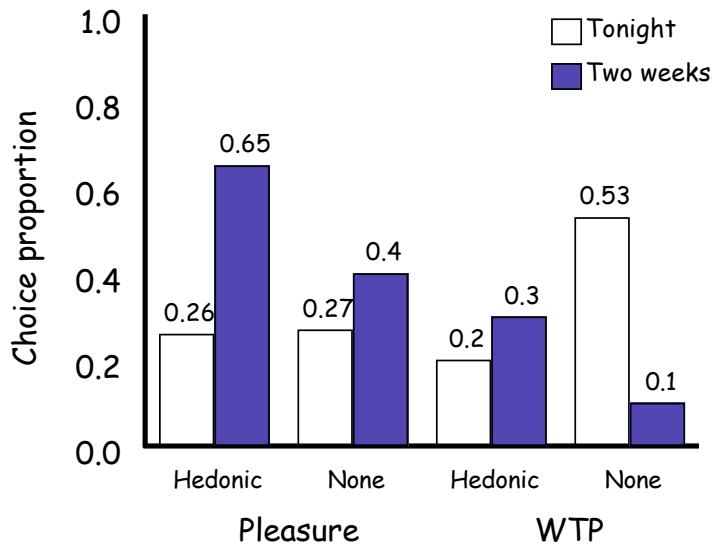


Figure 5

