Journal of Behavioral Decision Making
J. Behav. Dec. Making (2010)
Published online in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/bdm.713

The Use of Multiple Reference Points in Risky Decision Making

GREGORY J. KOOP* and JOSEPH G. JOHNSON
Miami University, Ohio, USA

ABSTRACT

Among psychologists and economists, prospect theory continues to be one of the most popular models of decision making. The theory's key property is reference dependence; specifically, how an individual's perception of loss or gain is dependent upon their starting point (i.e., the status quo). Although prospect theory is widely accepted, other authors have sought the inclusion of reference points besides the status quo. Initially these extensions focused on the importance of single reference points such as goals. More recently, authors have explained choice data by including multiple reference points within the value function. Multiple-reference-point theories generally assume that many choice situations possess an implicit or explicit goal, or point an individual will strive to obtain, and/or a minimum requirement (i.e., a ''lower bound'') above which an individual will strive to stay. In two experiments, we present evidence that individuals can utilize the minimum requirement, status quo, and goal within a single risky decision task. Participants most often chose to maximize their chance of reaching reference points even when that decision was riskier, resulted in lower expected value, resulted in lower expected utility, or ran contrary to the predictions of prospect theory. Furthermore, salience and uncertainty moderated the use of goals and minimum requirements as reference points. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS goals; status quo; prospect theory; gamble; uncertainty; reference points

INTRODUCTION

A great deal of human perception and evaluation is inherently comparative; that is, we are often inclined to view things not in absolute terms, but relative to expectations, standards, or benchmarks. Decades of research in social comparison have examined the comparative nature of judgments about others as well as oneself (Festinger, 1954; Mussweiler, 2003), but these comparative processes can also influence more basic processes such as psychophysical judgments (Brown, 1953). In judgment and decision making as well, considerable research has addressed this basic tendency (e.g., Kahneman & Miller, 1986; Sanbonmatsu, Kardes, & Gibson, 1991; Tversky, 1977). The extant research has focused primarily on identifying the existence of these comparative effects, as opposed to absolute judgment. For example, Markowitz (1952) and

* Correspondence to: Gregory J. Koop, Miami University, OH, USA. E-mail: koopgj@muohio.edu

prospect theory (Kahneman & Tversky, 1979) stressed the calculation of utility based on changes in wealth rather than absolute or final asset positions. However, much less research has been devoted to understanding the exact nature of the comparative benchmark, or whether multiple benchmarks may be relevant in a given decision situation.

In utility theories such as prospect theory, a basic assumption is that there is a single, fixed reference point; one often defined by the current level of wealth, or status quo. Although Kahneman and Tversky (1979: p. 286) acknowledge that ‘‘there are situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo,’’ relatively little empirical work has examined these circumstances (notable exceptions will be discussed in the following section). The current research focuses on a situation where there are three salient reference points that are believed to qualitatively represent the most pervasive reference points for comparative judgment in many situations. We show that individuals, in a single decision setting, seem to be systematically sensitive to multiple reference points when making decisions.

Evidence and theory for multiple reference points

Although the status quo may be the most commonly occurring evaluative benchmark, a number of other studies have shown that reference points besides the status quo (SQ) can have significant impacts on behavior. Specifically, research has examined the impact of goals (G) and aspirations as reference points (e.g., Heath, Larrick, & Wu, 1999; Lopes & Oden, 1999; March & Shapira, 1992; Sullivan & Kida, 1995). Like the status quo in prospect theory, goals can divide outcomes into regions of success and failure, gain and loss (Heath et al., 1999). Although aspirations are the most prominent example of reference points besides the status quo, survival requirements have also been mentioned as having an important impact on behavior (Lopes & Oden, 1999; March, 1988). Minimum requirements (MR) may refer to the minimum amount of income a worker needs to pay a month’s rent, or a ‘‘bottom line’’ required for an individual or corporation to remain financially solvent.

Early theories on multiple reference points suggested people combined these points into a single composite point (e.g., Olson, Roese, & Zanna, 1996; Ordóñez, 1998; Tryon, 1994) but a more recent study on perceptions of fairness has shown that people can simultaneously consider multiple reference points in their value judgments (Ordóñez, Connolly, & Coughlin, 2000). For instance, even while people can feel positively relative to one reference point like the SQ, they may feel negatively relative to an aspiration (Ordóñez et al., 2000). The asymmetries in subjective value that prospect theory often attributes to the SQ (i.e., ‘‘losses loom larger than gains’’) also appear around these other reference points.[1]

When individuals utilize multiple reference points, unique patterns of risk related behavior emerge. In performance assessment, investment managers have been shown to consider their current level of performance and a target performance rate (SQ and G; Sullivan & Kida, 1995). Contrary to the tenets of prospect theory, these investment managers were not universally risk averse when their performance was above their SQ. Managers were risk averse when the possibility of falling back to the SQ existed, but when they were guaranteed to stay above the SQ, they tended to be risk seeking in order to reach their G. Sullivan and Kida hypothesized that when multiple reference points are important for assessment, each of these reference points will concurrently impact behavior.

Additional evidence from animal behavior supports the claim that decision agents use multiple reference points. In particular, animal foraging behavior has shown that prospect theory’s assertion of uniform risk seeking for losses and risk aversion for gains is not accurate when one considers multiple reference points. Animal foraging theory predicts that animals will consider both a starvation threshold and a reproductive

[1]Although recent work by Jeffrey, Onay, and Larrick (2009) suggests the subjective experience involved in crossing the G may actually elicit more nuanced risk taking strategies. The authors suggest the presence of a ‘‘cushion effect’’ where individuals will again be more risk seeking when all outcomes surpass a salient goal.

DOI: 10.1002/bdm

energy threshold when making decisions about where to search for food (MR and G; Hurly, 2003). These two thresholds moderate risk-taking behavior in opposite directions. When just above the starvation threshold, animals were risk averse in order to avoid death whereas animals below that threshold were risk seeking in order to survive (Kacelnik & Bateson, 1997). Similarly, around the reproductive threshold, animals were risk seeking in order to surpass this biologically important reference point (Hurly, 2003).

In sum, substantial empirical evidence suggests that both human and nonhuman decision makers may employ reference points besides just the SQ. Furthermore, it may not be just that these reference points are considered *instead*, but that they are considered *in addition*. Thus, although the calculations of prospect theory allow for any reference point to be used, they do not allow for multiple reference points to each affect preferences within the same choice context. There have been, however, theories that do allow for the joint inclusion of multiple reference points.

Variable risk preference models

March and Shapira (1992). The variable risk preference models of March and Shapira (1988, 1992) explicitly acknowledge the potential impact of both the MR and G. Their model provides an interesting alternative to holistic, utility-based theories by describing how pursuit of risky courses of action depends on attentional mechanisms. That is, it models how a decision maker's attention may shift between a focus on the MR and G, and predicts how behavior may be impacted as a result. However, March and Shapira (1992: p. 175) assume an individual ''attends either to the aspiration level or to the survival point but not to both.'' In other words, although their theory allows for both MR and G to affect behavior, they would seem to do so in a mutually exclusive way.

Security-potential/aspiration theory

Lola Lopes' (1987) theory also models a decision maker's differential attention to surviving and thriving, which she describes as being ''security-minded'' or ''potential-minded.'' In terms of an expected utility calculation, these foci result in additional decision weight being placed on lower or higher outcomes, respectively. Although this treatment acknowledges the importance of these motivational forces, it also does not explicitly involve the notion of reference points (although it allows for coding of gains and losses more generally). For example, although a security-minded individual may pay more attention to low outcomes, she or he may not distinguish between low outcomes that are above versus below a MR threshold. Lopes' theory does explicitly include the notion of goal pursuit in a second component to her theory. Specifically, she assumes that choice options are also evaluated based on their probability of reaching a G. This second component is integrated with the holistic utility calculation to provide an overall assessment of each option.

Decision affect theory

The decision affect theory of Mellers, Schwartz, Ho, and Ritov (1997) acknowledges that outcomes are not only evaluated relative to the SQ as per prospect theory, but necessarily involve expectations. Mellers et al. (1997) suggest that emotions from counterfactual comparisons are crucial to decision making (see also work on regret theory: Bell, 1982; Loomes & Sugden, 1982; Zeelenberg, 1999). For example, even when an individual receives a positive outcome, the outcome may evoke negative feelings if an alternative outcome was better. In other words, ''gains can be disappointing, and losses can be elating'' (Mellers et al., 1997: p. 427). Consequently, options are not merely evaluated relative to a SQ, but involve expectations and lead to choices based on expected feelings. The importance of expectations in subjective evaluation parallels subjective changes incurred while crossing reference points. However, decision affect theory does not explicitly incorporate these overarching reference points (such as long term goals), instead focusing these evaluations on the contemporaneous alternative options.

DOI: 10.1002/bdm

Tri-reference-point theory

Wang and Johnson (2009; see also Wang, 2008 for a brief overview) introduced a tri-reference point (TRP) theory of risky decision making that explicitly considers the effects of the three reference points: MR, SQ, and G. TRP theory makes several specific assumptions about reference dependence (see Wang & Johnson, 2009, for motivations and full theoretical and mathematical treatment). First, it assumes that decision makers simultaneously desire to surpass a G, stay above a MR, and improve from their SQ. These reference points effectively carve the outcome space into distinct regions of failure (below MR), loss (at or above MR but below SQ), gain (between SQ and G), and success (at or above G). These regions may each have specific value functions, although it is typically assumed that these are parameterized jointly (e.g., they have identical curvature). Additionally, TRP theory assumes that there are psychological benefits to reaching the G, and detriments to falling below the MR, that produce (dis)utility that is not necessarily contingent upon increased or decreased objective outcomes. For example, if a salesperson reaches her sales target (G), she may indeed receive a monetary bonus that increases her subjective value, but also satisfaction, praise, job security, etc. that would produce additional utility. However, if a worker earns money below his sustenance budget (MR), the associated subjective value is greatly diminished even if the amount or utility of the earned money is not directly affected.

TRP theory thus predicts that a constant increase in value will be subjectively more impactful when it results in crossing a reference point into a different outcome region. For example, an increase in $50 will be more meaningful if it changes failure into "mere" loss, or "mere" gain into success, compared to when it "only" increases the magnitude of a loss or gain. Mathematically, this can simply be modeled by upward or downward shifts in the value function that produce different limits from above and below and thus a step discontinuity (see Wang & Johnson, 2009). Alternatively, continuity could be maintained by using multipliers such as prospect theory's λ, or appropriate piecewise sigmoidal functions. In any case, this means a key assumption of TRP is that the value function will be steepest in the neighborhood of the reference points.

TRP thus draws a number of testable contrasts with prospect theory. Unlike prospect theory, TRP allows for this increased steepness (slope) at three distinct regions across the value continuum, rather than solely at (subjective) zero corresponding to the SQ. Furthermore, regardless of which side of the SQ a decision maker is on, TRP assumes decisions will be moderated by the possibility of crossing other reference points. Instead of being universally risk averse for gains or risk seeking for losses as predicted by prospect theory, TRP predicts individuals will show strong risk-seeking tendencies when the riskier option allows them the opportunity to move above a reference point. Similarly, people will be risk averse when the safer option guarantees them the opportunity to remain above the reference point. Consider a choice between gamble A {450, .5; 550, .5} and gamble B {400, .5; 600, .5} where the MR has been set at 0, the SQ at 300, and the G at 600.[2] Because gambles A and B are entirely above the SQ, prospect theory predicts choice of A because it is the less risky (variable) option. In contrast, TRP predicts choice of B because it allows the individual to reach the G of 600.

These divergent predictions result from differences in how diminishing marginal utility is treated in the two theories. Instead of continually decreasing as one moves away in either direction from the SQ (as in prospect theory), TRP predicts that marginal differences will increase again at the G (as x increases) and the MR (as x decreases) due to the value function shifts. Furthermore, another assumption of TRP suggests differences in the size of the marginal differences around these reference points—the largest change is predicted to be at the MR, followed by the G, and finally, the SQ. This assumption is based on evolutionary theory (for full theoretical treatment, see Wang, 2002 or Wang, 2008) and fits with the idea of "security-first"

---

[2]We use the notation $\{x, p; y, q\}$ to represent a standard gamble with p probability of winning x, and q probability of winning y.

DOI: 10.1002/bdm

in business management (e.g., Roy, 1952). Because the MR is the most important, it should induce the most reference-point-dependent behavior according to TRP, followed by the G and the SQ.

Novel empirical tests

There are a number of theories in behavioral decision making that either implicitly or explicitly acknowledge three key reference points in comparative judgment: MR, SQ, and G. However, these theories differ with respect to exactly which reference points are important, and how their effects are manifest. To the best of our knowledge, however, no empirical test has directly manipulated all three reference points to see which one(s) affect behavior, and to what degree (e.g., Sullivan & Kida, 1995, employed only SQ and G). The primary goal of the current research is to identify whether all three reference points can indeed impact choice behavior. Consequently, a secondary goal is to support the theoretical stance (in the broadest sense) offered by multiple-reference-point (mRP) theories and show that the inclusion of multiple reference points can increase descriptive power. Similar to Heath et al.'s (1999) claim that goals can alter choice in a fashion consistent with PT, we suggest that goals *and* minimum requirements (in addition to the status quo) can alter choice behavior within a single choice context. Although a wide body of literature offers theoretical support to mRP-dependent behaviors unaccounted for by prospect theory, no one has specifically sought to empirically test the predictions of an mRP account.

Accordingly, the studies presented here induce three reference points within a gambling task and show that these points moderate individuals' risk strategies. Furthermore, the studies show that individuals will choose gambles that allow them to cross each of these reference points even when in direct competition with more lucrative gambles (higher expected value; Experiment 1), or safer gambles (less variable; Experiment 2) thereby showing choice behavior in accord with predictions of mRP-dependence. In order to definitively compare mRP accounts and prospect theory or expected utility explanations for behavior, Experiment 2 also includes stimuli that are diagnostic among strategies. For example, one gamble might have higher expected utility whereas the other gamble, although lower in expected utility, allows the participant to cross a reference point. Finally, some theories suggest even ambiguous or uncertain minimum requirements (Borch, 1968), goals (Bordley & LiCalzi, 2000; Oden & Lopes, 1997), or both (Wang & Johnson, 2009) can affect behavior, yet these effects have not yet been empirically examined. To the extent that reference points truly impact behavior, then adding uncertainty to the existence or location of these reference points should result in behavioral changes as well (presumably by decreased influence of the reference point). Indeed, our data from Experiment 1 show that if reference points become uncertain, their behavioral impact will decrease relative to unambiguous reference points.

EXPERIMENT 1

Participants

College students ($N = 155$) from introductory psychology classes at a Midwestern university participated in this study. Participants selected the experiment through an online sign-up site that allowed them to choose between many different experiments. For their involvement, participants received course credit; additionally, they were paid between \$1 and \$13 dollars (mode of \$5) based on their decisions, as described below.

Design and stimuli

The three reference points functionally divide the outcome space X (where $x \in X$) into failure ($x < \text{MR}$), loss ($\text{MR} \leq x < \text{SQ}$), gain ($\text{SQ} < x < \text{G}$), and success ($\text{G} \leq x$).[3] Let $\text{A} = a_1 a_2$ represent a single binary gamble A

[3]We omit here the possible outcome $x = \text{SQ}$, which would simply denote maintenance of the status quo.

DOI: 10.1002/bdm

Table 1. Gamble pairs in Experiment 1

Pair number	Gamble A		Gamble B		Common outcome (a_1, b_1)	Reference point involved (a_2, b_2)
	a_1	a_2	b_1	b_2		
1	940	960	580	1220	Failure	MR
2	1880	920	1600	1100	Loss	MR
3	3840	760	3300	1200	Gain	MR
4	4900	600	4220	1180	Success	MR
5	880	1920	620	2080	Failure	SQ
6	1720	1880	1420	2080	Loss	SQ
7	3700	1800	2820	2580	Gain	SQ
8	4600	1800	4200	2100	Success	SQ
9	760	3860	400	4120	Failure	G
10	1800	3700	1160	4240	Loss	G
11	3100	3600	2360	4240	Gain	G
12	5060	3680	4400	4240	Success	G

Note: Gamble pairs in Experiment 1 with MR = 1000 lira, SQ = 2000 lira, and G = 4000 lira. "Common outcome" is the region shared by a_1 and b_1. "Reference point involved" is the reference point that is straddled by a_2 and b_2—that is, a_2 is below this reference point and b_2 is above it. Exchange rate is $1 to 400 lira.

with outcomes a_1 and a_2 occurring with probability $\Pr(a_1) = \Pr(a_2) = p = .5$. Pairs of binary gambles were created such that, for each pair {A, B}, a_1 and b_1 were in the same region, whereas a_2 and b_2 were in adjacent regions—that is, the second outcomes straddled a reference point (see Table 1). For example, in Pair 9, both a_1 and b_1 represent failure, whereas a_2 represents a gain and b_2 represents a success—these outcomes straddle the G. Holding constant the reference point straddled by a_2 and b_2 and moving a_1 and b_1 across all functional regions produces four gamble pairs. Repeating this for each of the three reference points produces a total of $4 \times 3 = 12$ gamble pairs seen by each participant. The values of a_1, b_1, a_2, and b_2 were chosen so that *A* always had a higher expected value (presented as a difference of 50 lira, which was 12.5 cents in real money), whereas B always had the better functional outcome (i.e., allowed participants to cross a reference point).

This study intends to demonstrate that multiple reference points affect behavior. If this is the case, then one would expect that weakening the manipulation would also weaken the effect. To this end, we decided to add uncertainty to the reference points as a way to potentially moderate their influence (Borch, 1968; Bordley & LiCalzi, 2000; Oden & Lopes, 1997; Wang & Johnson, 2009). Specifically, we manipulated the reference point type between-subjects using a strong (certain) condition and a weak (uncertain) condition. In the strong condition ($N = 72$), reference points MR, SQ, and G were well-defined and revealed to participants. In the weak condition ($N = 83$), SQ was held fixed but MR and G were expressed in terms of symmetric probability distributions. The mean of each distribution was set equal to the associated value from the strong condition.

Procedure

We conducted the experiment in a computer lab with participants seated at nonadjacent computers. After all participants were seated at their computers, an experimenter introduced the gambling task and notified the participants that their choices would determine their winnings in real money to be paid at the conclusion of the experiment. Additionally, the amount of their winnings would determine the number of entries into a drawing for a mystery prize worth approximately $20. During this introduction, the experimenter held a wrapped gift box and placed it in a location where it was visible to all participants to enhance salience. Participants then read through instruction slides at their own pace to learn the final details of the task.

DOI: 10.1002/bdm

We wanted the gamble outcome values to have maximum impact on participants, so we sought to nominally introduce more variability in these values and eliminate presentation of decimal values that promote rounding. To accomplish this, dollar amounts were represented in Italian lira with an exchange rate of \$1 to 400 lira. Gamble payouts ranged from 20 lira to 5060 lira. Participants were made fully aware that their payout would be determined by the gamble they selected on a randomly chosen trial and that the outcome of that chosen gamble would also be randomly determined. Participants were told that they earned \$5 for showing up at the experiment but could trade those \$5 for the right to play whichever gamble was randomly selected from those they were about to see. This was intended to establish a SQ of \$5 (2000 lira). No other explicit mention of the SQ, or of comparing outcomes to the SQ, was made.

Inducing reference points: Bonus drawing

Just as Heath et al. (1999) emphasized the use of "mere" goals (for our purposes, goals that are not directly tied to a discontinuity in the primary monetary reward), we sought to implement G and MR that were not simply monetary bonuses. To establish multiple reference points, participants were given various earning benchmarks that determined their entries into the mystery prize drawing. Thus, gamble stimuli and immediate task payouts were represented in monetary terms, whereas reference points MR and G were induced in the currency of bonus drawing entries. The use of two separate currencies made it possible to isolate the reference points MR and G from the main task and prevent explicit incorporation of the reference points in recoding one's value function during the task. Explicit computation involving reference points was also inhibited due to the probabilistic nature of the bonus drawing (and the lack of knowledge about the number of competitors therein). In sum, we dissociated the impact of reference points and monetary reward on subjective value through two means: Temporal separation (falling below the MR or surpassing the G had no affect on immediate payout, but would only have an effect weeks later when the drawing would be held) and through incommensurate currencies (lira in the main task versus raffle tickets toward a mystery prize). This distinction was made clear to participants, such as by stressing (in the parlance of the task) that the achievement of reference points would have no bearing on the amount earned at the experiment's conclusion.

Although the SQ was held constant for all participants, the MR and the G induction differed between two experimental conditions in order to manipulate the strength of the reference point effect. In the strong (certain) condition, the MR was set by requiring participants to earn at least 1000 lira to gain any entries in the bonus drawing. After surpassing 1000 lira, the number of entries earned for the bonus drawing was calculated by dividing the amount of lira they earned by 100 (e.g., 2000 lira = 20 entries). If participants met or surpassed the G of 4000 lira, their entries into the bonus drawing would be doubled (e.g., 4000 lira = 80 entries).

In the weak (uncertain) condition, rather than having a set MR of 1000 and G of 4000, participants were told that these reference points would be randomly selected following a symmetric probability distribution with 1000 lira and 4000 lira as the means (Figure 1). In the weak condition, the MR had a 38% chance of being between 900 and 1100 lira with other possible values ranging from 500 lira to 1500 lira. The G also had a symmetric distribution with possible values ranging from 3500 lira to 4500 lira. The number of entries earned in the bonus drawing was actually calculated using the same equation as in the strong condition.

Manipulation check

In order to ensure participants understood how to earn entries into the bonus drawing, the instruction slides were followed by a short quiz that asked participants to recall the reference point values and calculate the number of bonus drawings associated with two different scenarios (see Appendix A, B). The experimenter corrected any mistakes and further explained the system to each participant before beginning the main task. Additionally, following each block of 12 trials, participants answered questions about the values (or possible

DOI: 10.1002/bdm

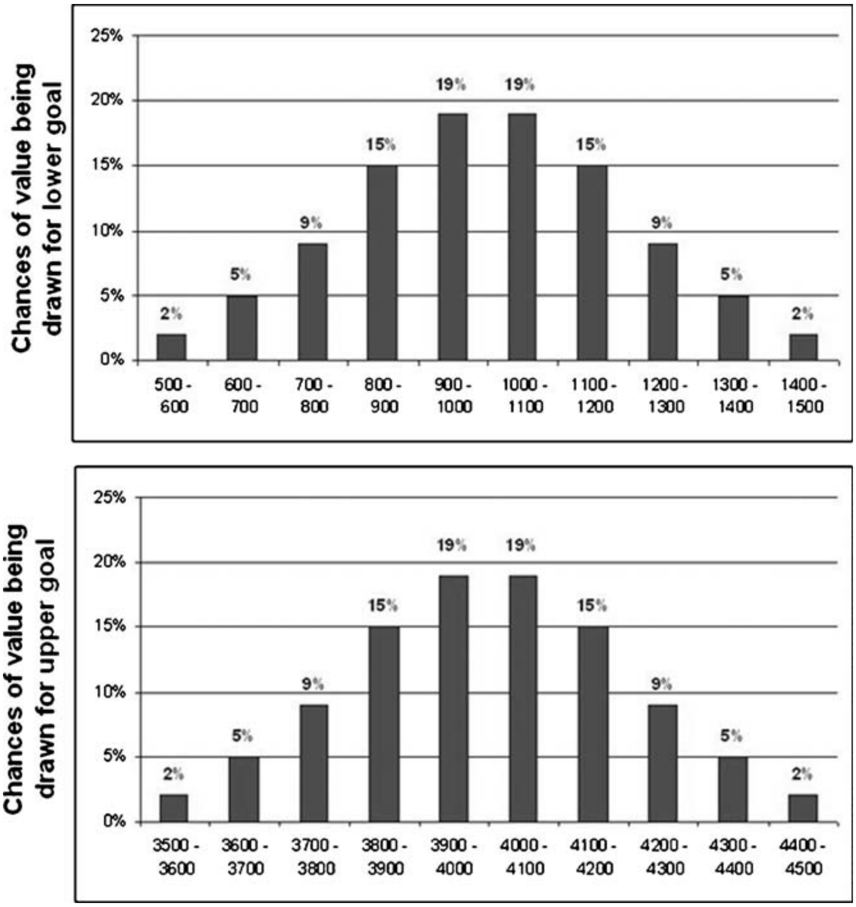

Figure 1. The distributions for possible values of MR and G in the weak condition. These charts were presented to participants during the instructions

values) of the reference points. We required perfect performance on the quiz to continue to the next block. If a participant failed this quiz three consecutive times during any one presentation (between any two blocks), the computer displayed the correct answers and returned the participant to the previous block of gambles. In these instances, data from the initial attempt on that block were discarded and only the new responses were used in analyses.

Gambling task

The gambling task consisted of three identical blocks of the 12 gamble pairings (presented via E-prime software; see Figure 2). The presentation order of gamble pairs was separately randomized for each block of each individual. Each of the gamble pairs had three choice options: An ‘‘A’’ gamble, a ‘‘B’’ gamble, and ‘‘Indifferent.’’[4] The presentation of multiple identical blocks allowed for the calculation of choice proportions with precision of 0.33, rather than single choice estimates.

[4]We allowed participants the option of choosing ‘‘indifferent’’ in order to prevent truly random choices from polluting genuine choice strategies.

DOI: 10.1002/bdm

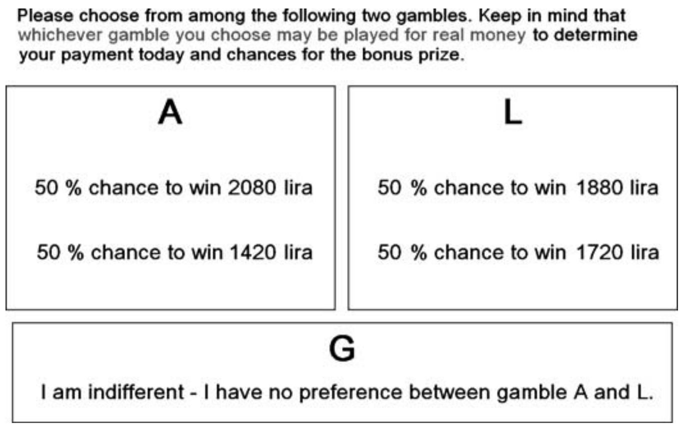

Figure 2. A screen shot from the gambling task. Participants were asked to select their preferred gamble by pressing ‘‘A’’ or ‘‘L.’’ Participants could also signify indifference by pressing ‘‘G’’

Payment

After completing the instructions and gambling task, participants went to another room in order to receive their payment following a procedure that was made abundantly clear during the initial instructions. Upon arrival, a single trial from the completed gambling task was randomly chosen. The participant’s chosen gamble on that trial was determined from their data and presented to the participant who then had the option of playing that gamble or leaving with the initial $5.[5] If the participant chose to play the gamble, the outcome was randomly selected. Any random selection was transparently done using random number lists, to prevent participants from being suspicious or doubting the statistical nature of the gambles (this was explained in detail in the instructions). The experimenter then paid each participant according to the gamble outcome and calculated the number of entries earned for the bonus drawing, which was performed at a later date. Bonus drawing winners were notified via e-mail and paid $20 in cash instead of a mystery prize of that value.

Results

Recall that the B gamble always had the better functional outcome in terms of the reference points (bonus drawing) but also had a lower expected value. During the experiment the left/right presentation order of A and B gambles was counterbalanced without statistically significant order effects. For clarity, however, let B hereafter represent the gamble predicted by mRP-dependence—that is, the gamble predicted for a decision maker with the primary motivation of achieving the three reference points (rather than maximizing expected value). Thus, the dependent variable measuring the extent to which participants utilized each reference point (MR, SQ, G) was calculated by the choice frequency for B gambles excluding trials with an indifferent response: Pr(B|∼indifferent). These individual proportions were then averaged across subjects. Two participants from the strong condition were excluded from all analyses due to an unusually high number of indifferent responses (34/36 and 36/36). The null hypothesis of insensitivity to reference points suggests a choice proportion of 0.50; this value was used in one-sample *t*-tests reported below. Note this is more conservative (more difficult to claim mRP-dependence) than a null hypothesis of expected value

[5]If the participant selected the indifference response on the trial selected for determining payment, then one of the two gambles was randomly selected. Participants were told this in the instructions.

DOI: 10.1002/bdm

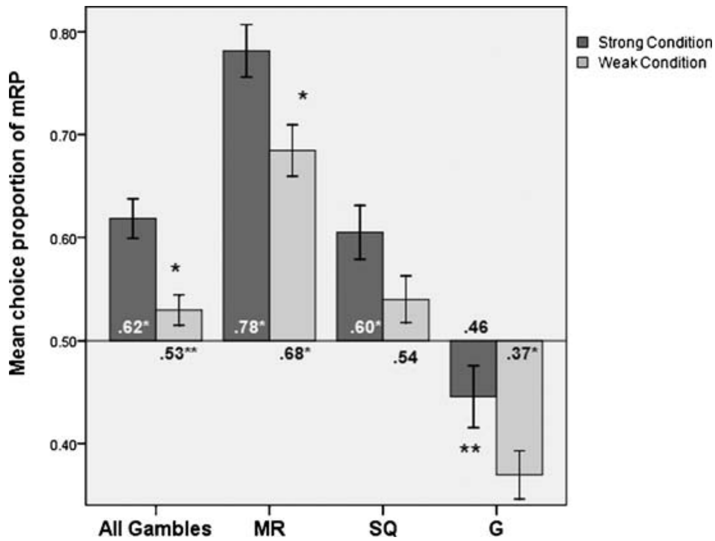

Figure 3. Mean (across participants) conditional proportion of choosing the mRP-predicted gamble (±SE) in Experiment 1, shown for all gambles and for each of the reference points separately. Where indicated, choice proportions differed from chance, $^{*}p < .01$, or $^{**}p < .05$. Additionally, choice proportions differed between conditions everywhere except around the SQ

maximization, which would predict a choice proportion of Pr(B|∼indifferent) = 0.00. The aggregate results are presented in Figure 3, and the choice proportions for individual gambles are presented in Table 2.

In the strong condition, participants showed a general tendency to choose the gambles predicted by mRP-dependence more frequently than predicted by chance, $t(70) = 6.15, p < .01$. Across all reference points (12 pairs), participants chose in line with mRP-dependence 61.83% of the time. Analyses were also conducted separately for the four pairs associated with each reference point—that is, on those trials where the

Table 2. Predictions and results for gamble pairs in Experiment 1

	Risk		Expected value		Pr(B|∼indiff)		*p*	
Pair number	$\|a_1–a_2\|$	$\|b_1–b_2\|$	EV(A)	EV(B)	Strong	Weak	Strong	Weak
1	**20**	640	**950**	900	0.51	0.29	.85	**<.01**
2	960	**500**	**1400**	1350	0.85	0.83	**<.01**	**<.01**
3	3080	**2100**	**2300**	2250	0.88	0.76	**<.01**	**<.01**
4	4300	**3040**	**2750**	2700	0.83	0.83	**<.01**	**<.01**
5	**1040**	1460	**1400**	1350	0.40	0.37	.06	**<.01**
6	**160**	660	**1800**	1750	0.47	0.40	.55	**.02**
7	1900	**240**	**2750**	2700	0.74	0.69	**<.01**	**<.01**
8	2800	**2100**	**3200**	3150	0.70	0.65	**<.01**	**<.01**
9	**3100**	3720	**2310**	2260	0.41	0.29	.09	**<.01**
10	**1900**	3080	**2750**	2700	0.48	0.38	.59	**<.01**
11	**500**	1880	**3350**	3300	0.44	0.40	.26	**.02**
12	1380	**160**	**4370**	4320	0.54	0.42	.45	.09

Note: Predictions and results for gamble pairs in Experiment 1. Because each choice outcome is equally probable, risk can be operationalized as the absolute difference between outcomes. Bold denotes the predictions of risk aversion and EV. Choice proportions are reported for each gamble pair in both the strong and weak conditions from Experiment 1. All *p* values are the result of one-sample *t* tests against a null choice proportion of 0.5.

DOI: 10.1002/bdm

divergent outcome regions for gambles A and B straddled a reference point (e.g., in Table 1, trials 1, 2, 3, and 4 are around the MR). Gambles around the MR showed the most behavior in line with mRP-dependence, with participants choosing the B gamble 78.15% of the time, statistically more than predicted by chance, $t(68) = 10.98$, $p < 01$. Participants chose in line with mRP-dependence for gambles around the SQ in 60.49% of cases, $t(69) = 4.00$, $p < .01$. Unlike the MR and SQ, preference in line with mRP-dependence around the G was not statistically different from chance, $t(69) = -1.80$, $p = .08$, as participants only chose the B gambles on 45.53% of gambles.

Across all reference points in the weak condition (12 pairs), participants chose the gamble in line with mRP-dependence 53% of the time, which was statistically more than predicted by chance, $t(82) = 2.02$, $p = .047$. Although participants in the weak condition tended to choose in line with mRP-dependence overall, among individual reference points these results were only replicated for the MR, where participants chose in line with mRP-dependence 68.45% of the time; this was more than predicted by chance, $t(82) = 7.38$, $p < .01$.

Previous research has suggested effects of reference point ambiguity on the MR and the G individually (Borch, 1968; Bordley & LiCalzi, 2000; Oden & Lopes, 1997), and collectively (Wang & Johnson, 2009). Indeed, the strong condition showed a higher proportion overall of mRP-dependent choice than the weak condition, $t(152) = 3.7$, $p < .01$. When these results are further dissected on individual reference points, the predictions made by mRP-dependence again fit the data well. Participants in the strong condition showed a stronger preference for the gambles in line with mRP-dependence than participants in the weak condition for gambles around the MR, $t(150) = 2.69$, $p < .01$, as well as gambles around the G, $t(151) = 2.02$, $p < .05$. Differences for gambles around the SQ between the strong condition and weak condition were not statistically significant, $t(151) = 1.89$, $p = .06$. However, this is not surprising, as the SQ did not change between conditions. All participants started with the same amount of money to begin the task thereby creating a fixed SQ, even in the weak condition that contained additional uncertainty.

Discussion

Despite showing an overall preference for the gambles predicted by mRP-dependence, the data from Experiment 1 did not entirely fit with our initial predictions. In general, participants failed to show mRP-dependent behavior on gambles around the G. That is, their choice of gambles that would allow them to reach the G was not any different than would be expected by chance. This prompted us to more closely examine our stimuli, which revealed a possible explanation: The gamble pairs that straddled the G were confounded by risk aversion. On three of the four gamble pairs around the G, the gamble predicted by mRP-dependence was also the riskier (more variable) option. The effect of facing the safer gambles for an immediate payout could have excessively decreased preference for the gambles that ‘‘merely’’ offered better chances at a bonus drawing. Also, on gambles around the MR, those that guaranteed values above the MR were safer (less variable) on three of four pairings, which may have inflated apparent mRP-dependent preference for these gambles.

It is important to note, however, that the reference points showed more impact in the strong condition than in the weak condition. Even if risk aversion drove preference for gambles around the MR and G, this risk tendency should have been the same across both strong and weak conditions. However, the decrease in mRP-dependent behavior when the reference points are weakened by uncertainty provides additional evidence that characteristics of the reference points themselves impact behavior, above and beyond risk attitude. In fact, when the reference point manipulation was weakened, choices significantly favored the gamble with higher expected value, not the achievement of unique reference points, for half of the gambles when considered individually ($p < .05$; see Table 2). Still, we created our stimuli to make mRP-dependent choice contradict choices based on maximized expected value, and although comparisons between conditions allow us to look at the global impact of reference points, we cannot fully assess the use of mRP-dependence relative to other strategies within Experiment 1. Specifically, we cannot directly compare mRP-dependent

DOI: 10.1002/bdm

predictions with strategies such as ‘‘pure’’ constant risk aversion, expected utility, or prospect theory within Experiment 1. Furthermore, just as weakening the reference points through uncertainty decreased their apparent utilization, perhaps increasing their salience would lead to increased usage.

The design limitation of this experiment regarding risk attitude, in conjunction with the realization that more salient reference points might increase their usage, motivated us to design a second experiment with two key design improvements. First, by controlling when gambles predicted by mRP-dependence lined up with lower-variance gambles, those predicted by expected utility, and those predicted by prospect theory, it became possible to measure explicitly whether people made decisions consistent with each of these strategies. Additionally, by making the reward scheme simpler and thus making the impact of crossing reference points clearer, it was hypothesized that more mRP-dependent behavior would emerge. In a second experiment with a revised drawing entry scheme and prize presentation reported below, we show an increase in behavior predicted by mRP-dependence relative to risk aversion, expected utility, and prospect theory.

EXPERIMENT 2

Participants

College students ($N = 55$) from introductory psychology classes at a Midwestern university participated in this task. Selection procedures and course-credit compensation were identical to Experiment 1.

Design and stimuli

The division of the outcome space was identical to Experiment 1. Once again gamble pairs {A, B} held the region of a_1 and b_1 in common while a_2 and b_2 straddled a reference point. This was duplicated over all three reference points creating a set of 12 gambles. Unlike Experiment 1, the values of a_1, b_1, a_2, and b_2 (Table 3) were chosen specifically to pit predictions made by mRP-dependence, risk aversion (RA), expected utility (EU), and prospect theory (PT) against one another on a set number of gambles. We follow convention in defining risk aversion as always selecting the gamble with the lower mathematical variability. Expected utility was formalized with a power utility function with an exponent of 0.8. Prospect theory was implemented using the equations of Tversky and Kahneman (1992) and with parameters $\alpha = \beta = 0.8$ and $\lambda = 2.25$.[6] For each reference point, the gamble predicted by mRP-dependence (again, arbitrarily the B gamble) was paired twice with a gamble predicted by EU and twice with a gamble predicted by risk aversion (Table 4). That is, on two gamble pairs the theories made divergent predictions and on two gamble pairs the theories made identical predictions. Additionally, on five gambles prospect theory and mRP-dependence also made divergent predictions, allowing for comparison of the two accounts (i.e., whether the G and the MR were important above and beyond prospect theory’s SQ).

Procedure and inducing reference points

The overall procedure for Experiment 2 was similar to Experiment 1 but incorporated some minor improvements. All reference points were well-defined (as in the strong condition from Experiment 1) and revealed to participants prior to beginning the gambling task. The values of the MR and G were identical to the strong condition of Experiment 1. The SQ was changed to 2500 lira in order to center it between the MR (1000 lira) and the G (4000 lira) making the exchange rate $1 = 500 lira.

[6]The parameter values are also similar to those in Tversky and Kahneman (1992) where $\alpha = \beta = 0.88$ and $\lambda = 2.25$; these parameter values vary somewhat in the extant literature.

DOI: 10.1002/bdm

Table 3. Gamble pairs in Experiment 2

Pair number	Gamble A a_1	Gamble A a_2	Gamble B b_1	Gamble B b_2	Common outcome (A1 B1)	Reference point involved (A2 B2)
1	940	960	580	1220	Failure	MR
2	2000	920	1720	1100	Loss	MR
3	3650	750	3300	1200	Gain	MR
4	4050	950	4300	1100	Success	MR
5	880	1920	720	2680	Failure	SQ
6	1900	2200	1420	2580	Loss	SQ
7	3850	1800	2970	2580	Gain	SQ
8	4600	2200	4200	2700	Success	SQ
9	640	3980	720	4000	Failure	G
10	1800	3700	1360	4240	Loss	G
11	3300	3860	2560	4500	Gain	G
12	5060	3640	4400	4200	Success	G

Note: Gamble pairs in Experiment 2 with MR = 1000 lira, SQ = 2500 lira, and G = 4000 lira. "Common outcome" is the region shared by a_1 and b_1. "Reference point involved" is the reference point that is straddled by a_2 and b_2—that is, a_2 is below this reference point and b_2 is above it. Pairs 4 and 9 included dominant gambles and were not included in analyses. Exchange rate is $1 to 500 lira.

Experiment 2 still used a bonus drawing to isolate reference points MR and G from the value function of the main task, but the reward and the entry scheme for the drawing were simplified. Rather than an experimenter presenting a mystery prize prior to the task, participants were informed via instruction slides that the prize for the bonus drawing was a $20 gift card to a list of possible stores, from which they could choose. Again, it was made clear that earning entries into this drawing had no effect on the immediate monetary payout following the task. To maximize the salience of the prizes, participants entered their gift card store choice immediately prior to beginning the gambling task.

In Experiment 1, we were concerned that the mathematical computations required to calculate drawing entries overtaxed some participants. Ability to perform these calculations was not of importance to our design or hypotheses; rather we simply wanted participants to be aware that the MR served as a lower threshold for entering the bonus drawing, and reaching the G increased their chances in the drawing. Therefore, we simplified the method by which entries were calculated. Participants were told that they had to exceed 1000 lira in order to earn any entries in the drawing. If they surpassed 1000 lira they would earn 5 entries into the bonus drawing and if they surpassed 4000 lira they would earn 10 entries in the drawing. Other than those reference points, the amount of entries in the drawing was not related to amount earned. The manipulation checks, gambling task, and payment procedure were identical to Experiment 1.

Table 4. Division of predictions in Experiment 2

Gamble A	Gamble B	Pair number (Table 3)
EU, RA	mRP	1, 6, 11
EU	mRP, RA	2, 7, 12
RA	mRP, EU	4, 5, 10
	mRP, EU, RA	3, 8, 9
PT	mRP	1, 2, 10, 11, 12

Note: The division of predictions made by mRP-dependence, risk aversion (RA), expected utility (EU), and prospect theory (PT) as parameterized in Experiment 2. The first four rows contain four gamble types with unique prediction combinations between mRP, RA, and EU, which are repeated across each of the three reference points. "Pair Number" refers to the gamble represented in Table 3. Across all gambles, mRP predicts 100% choice of B, RA and EU predict 50% choice of B, and chance also predicts 50% choice of B. In the subset of gambles directly comparing PT to mRP, the latter again predicts 100% choice of B whereas PT predicts 0% choice of B.

DOI: 10.1002/bdm

Table 5. Predictions and results for gamble pairs in Experiment 2

Pair number	Risk		EV		EU		PT		Results	
	$\lvert a_1–a_2 \rvert$	$\lvert b_1–b_2 \rvert$	EV(A)	EV(B)	EU(A)	EU(B)	PT(A)	PT(B)	P(B|∼indiff)	*p*
1	**20**	640	**950**	900	**241.09**	228.48	**−802.50**	−820.49	0.53	.68
2	1080	**620**	**1460**	1410	**336.16**	329.36	**−569.76**	−601.52	0.84	**<.01**
3	2900	**2100**	2200	**2250**	453.61	**471.74**	−301.71	**−243.52**	0.93	**<.01**
4	**3100**	3200	2500	**2700**	505.08	**538.95**	−222.92	**−168.88**	0.95	**<.01**
5	**1040**	1960	1400	**1700**	325.03	**372.93**	−598.45	**−416.36**	0.78	**<.01**
6	**300**	1160	**2050**	2000	**445.88**	434.35	−295.65	**−283.88**	0.37	**.01**
7	2050	**390**	**2825**	2775	**570.26**	568.12	−52.76	**85.30**	0.74	**<.01**
8	2400	**1500**	3400	**3450**	661.77	**673.90**	119.52	**226.67**	0.65	**<.01**
9	3340	**3280**	2310	**2360**	467.09	**477.30**	−292.40	**−274.50**	0.97	**<.01**
10	**1900**	2880	2750	**2800**	558.71	**559.52**	**−67.12**	−118.20	0.61	.53
11	**560**	1940	**3580**	3530	**696.46**	684.77	**265.68**	231.90	0.53	.6
12	1420	**200**	**4350**	4300	**812.57**	806.78	**405.89**	401.89	0.67	**<.01**

Note: Predictions and results for gamble pairs in Experiment 2. EV = expected value; EU = expected utility with x^{α} and $\alpha = 0.8$; PT = prospect theory with $u(x) = x^{\alpha}$ for $x > \text{SQ}$ and $u(x) = -(\lambda(\text{SQ-}x)^{\beta})$ for $x < \text{SQ}$, $\alpha = \beta = 0.8$ and $\lambda = 2.25$. Bold denotes prediction from corresponding strategy; mRP-dependence always predicts choice of B. Gamble pairs 4 and 9 had stochastically dominant options and were excluded from analyses. Results show choice proportions of mRP-predicted gambles and the *p* value resulting from one-sample *t*-tests against a null choice proportion of 0.5.

Results

Just as in Experiment 1, the dependent variable was choice proportion of B gambles given non-indifference—that is, Pr(B|∼indifferent). These individual proportions were again averaged across participants and analyzed across each reference point. As in Experiment 1, the null hypothesis of mRP-independence assumed a choice proportion of 0.50. Choice proportions for all gambles are presented in Table 5 and summarized in Figure 4.

Across all gamble pairs, participants showed a strong preference for the B gambles, $t(50) = 7.97$, $p < .01$. Participants chose the B gamble 67.3% of the time. Unlike Experiment 1, the general preference for gambles predicted by mRP-dependence was present when considering each and every reference point. Around the MR, participants chose the B gamble on 77% of pairings, significantly greater than chance, $t(50) = 9.01$, $p < .01$. Around the SQ, participants chose the B gamble on 64.13% of pairings, $t(50) = 6.70$, $p < .01$. Finally, on gambles around the G, participants chose the B gamble on 61.24% of gamble parings, $t(50) = 3.242$, $p < .01$.

Recall from Experiment 1 that picking the B gamble could have been due in some cases to the use of other strategies such as RA, EU, or PT. The current experiment made it possible to isolate the use of each of these individual strategies by creating subsets of gamble pairs in which RA, EU, and PT made divergent predictions from mRP-dependence. Out of the 12 unique gamble pairs (Table 3),[7] mRP-dependence made divergent predictions with RA on six pairs, with EU on six pairs, and with PT on five pairs. Three separate analyses were conducted, each involving only the pairs where mRP-dependence and the rival account made different predictions. For these analyses, the null hypothesis assumed a choice proportion of 0.50, which would indicate preference for neither strategy. Preference for the rival strategy would be revealed by always choosing A—in terms of our dependent variable, Pr(B|∼indifferent) = 0.00.

The data revealed behavior strongly in accord with mRP-dependence and contradictory to alternative explanations (see Figure 4). For only those gambles where RA predicted A and mRP-dependence predicted B, participants continued to choose the B gambles 63.65% of the time, which was greater than chance, $t(50) = 3.72$,

[7]Structuring the stimuli to place the different accounts in direct competition with each other imposed extra constraints that created two pairs in which there was a dominant option. These gambles were not included in analyses.

DOI: 10.1002/bdm

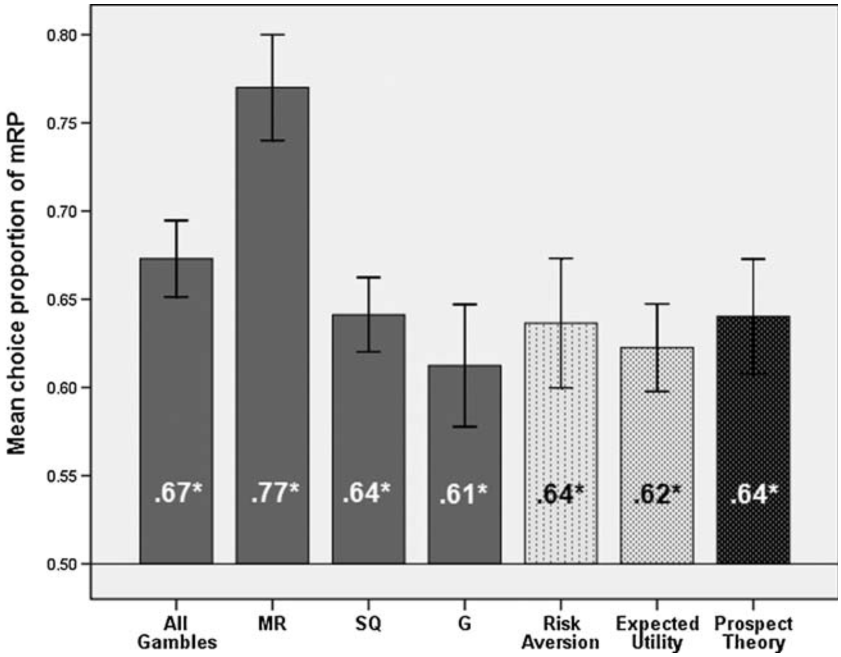

Figure 4. Mean (across participants) conditional proportion of choosing the mRP-predicted gamble (±SE) in Experiment 2, shown separately for each of the reference points and for gamble subsets used to diagnose strategy use. Strategy labels in the last three categories indicate gamble subsets where the labeled strategy predicted choice of the gamble opposite mRP-dependence. All choice proportions differed from chance, $p < .01$

$p < .01$. Similarly, for only those gambles where EU predicted A and mRP-dependence predicted B, 62.25% of choices were for B, which was greater than chance, $t(50) = 4.93$, $p < .01$. Finally, for those gambles where PT predicted choice of A, but mRP-dependence predicted choice of B, participants chose the latter 64.03% of the time, also greater than predicted by chance, $t(50) = 4.324$, $p < .01$. Note that these significance tests compare mRP-dependence to predictions of chance, which are much more conservative than tests against the specific theories (which each predict no choice of B, rather than 50% choice of B).[8]

General discussion

Our data show that three specific reference points (MR, SQ, and G) can collectively moderate human choice behavior within the same choice context. Although Kahneman and Tversky (1979) recognized the possibility of other reference points besides the SQ impacting behavior, that claim lacked direct empirical support. Some other instances of mRP-dependence in the literature assumed the impact of these reference points on behavior would be mutually exclusive. Experiments 1 and 2 systematically tested assumptions of no reference points (EV, EU, and RA) and a single reference point (prospect theory's SQ), and showed that the collective use of multiple reference points is the best explanation of the data.

[8]Our data analyses distinguish between multiple theories (RA, EU, PT and mRP), but of course every conceivable theory is not (and cannot be) included. It could be that some other characteristic of the stimuli, not formalized by any of these strategies, drove choice proportions away from the $\Pr(B|\sim\text{indifferent}) = 0.50$ expectation of chance. In order to address this possibility, 41 participants completed the stimuli used in Experiment 2 with instructions and a payoff scheme that were devoid of any mention of any reference points. Across all reference points, choice proportions of mRP-predicted gambles were lower in the control condition than in Experiment 2, $t(87.75) = 2.85$, $p < .01$. Differences between these participants and those in Experiment 2 were also significant when considering only the MR gambles, $t(90) = 2.04$, $p < .05$ or G gambles, $t(90) = 3.20$, $p < .01$, but not the SQ. These comparisons fit with our interpretations of reference-dependent choice when compared against a null hypothesis of random choice strategies.

DOI: 10.1002/bdm

Experiment 1 demonstrated the need to move beyond a strict interpretation of prospect theory by showing more nuanced risk-taking strategies around the MR and the SQ. Around these two reference points, participants were generally risk seeking when they had the possibility of surpassing a reference point and risk averse when in danger of falling below one. Additionally, Experiment 1 showed the impact of reference points by comparing choice proportions across a relatively strong certain-reference-point condition and a weaker uncertain-reference-point condition. Choices were more likely to follow these reference points in the strong condition, suggesting that characteristics of the reference points MR and G can jointly impact choice behavior. Experiment 2 placed the predictions of mRP-dependence, RA, EU, and PT in direct competition. With a simplified structure and more salient goals, Experiment 2 showed individuals strove to cross or stay above each of the three reference points. Contrary to prospect theory, these data showed that individuals constructed a more nuanced idea of failure and success and did not simply divide the outcome space along the SQ. Participants' choice behavior demonstrated an awareness of all three reference points. Consequently, the data were in accord with mRP-dependent theories rather than risk aversion, expected utility, or prospect theory.

In addition to showing participant use of multiple reference points, Experiment 2 also showed that the three reference points (MR, SQ, and G) had differential impacts on behavior. Heightened relative importance of the MR is predicted by at least one mRP-dependent theory (TRP) and fits with evolutionary theory and the business mantra of ''safety first'' (Wang, 2002). The SQ, being less important than the MR, suggests that although the general principles of prospect theory are sound, a strict adherence to SQ-dependent decreasing marginal utility may be oversimplified as an explanation of how people perceive and judge success. The idea that an individual may be confronted with multiple reference points when making a decision is not a new one. Although the behavioral decision making literature has acknowledged multiple reference points, much of this research assumed these reference points would be combined into a single, composite reference point that would behave similarly to the SQ (Olson, Roese, & Zanna, 1996; Ordóñez, 1998; Tryon, 1994), or that individuals attend to only one of these reference points at a time (Lopes & Oden, 1999; March & Shapira, 1992). In general, the body of research concerning aspirations and survival requirements has utilized multiple separate reference points (e.g., Heath et al., 1999; Lopes & Oden, 1999; March, 1988; March & Shapira, 1992). Lopes (1987), Higgins (1997) and others have also incorporated similar concepts as motivational influences, while Neale and Bazerman (1991) have used multiple reference points in terms of negotiations. Much as this literature has shown that individuals can judge success and failure differently relative to multiple reference points (Ordóñez, 1998), our findings show that the three specific reference points can also jointly impact choice behavior.

One mRP account, TRP (Wang & Johnson, 2009), assumes that in a choice context such as the one used in this series of experiments, individuals will *simultaneously* consider MR, SQ, and G throughout the experiment. Although our aggregate choice data seem to fit with this account, it is also possible that individuals demonstrated an adaptive use of reference points. That is, one could argue that they did not simultaneously consider all reference points on every trial, but only those that were relevant for each specific trial. Within the confines of the present experiment, it is impossible to disambiguate these two possibilities. Although self-report data showed that if participants considered the MR they were likely to have considered the G as well,[9] this is not sufficient to draw strong conclusions about the *simultaneous* use of the reference points. All we can say is that across the experiment as a whole, all three reference points seemed to be important. That the present experiments cannot directly assay simultaneous (within-trial) reference point use shows the need for future studies to critically examine these possible explanations.

Multiple theories (e.g., Lopes & Oden's SP/A, 1999; Wang & Johnson's TRP, 2009) critique prospect theory's ability to account for different choice strategies within a given realm—for example, individuals demonstrating both risk-averse and risk-seeking behavior in the realm of gains. Whereas the SP/A model calls for a break from prospect theory's value function (with the addition of a new aspiration criterion), TRP

[9]A statistically significant correlation from Experiment 1, $r(47) = .65$, $p < .01$.

DOI: 10.1002/bdm

suggests the need for an increasingly refined theory within the same framework. Future studies exploring divergent predictions between theories (e.g., holding multiple reference points simultaneously) will help determine which account best describes human choice behavior.

A possible critique of the current findings is that they simply reflect induced artificial reference points without any real-world counterpart. Fortunately, several real-world examples of imposed reference points support our methodology. Sullivan and Kida's (1995) study with investment bankers dealing with company mandated quotas produced similar findings. Their results showed variable risk-taking strategies within the realm of gains causing them to conclude "risky decision making may be more complex than models such as Prospect Theory would suggest in certain decision contexts" (Sullivan & Kida, 1995: p. 82). Their findings within the context of investment banking demonstrate that imposed multiple reference points are at least plausible in everyday settings.

More recently, Wang and Johnson (2009) utilized the standard corporate salary structure in Shanghai, China to implement the predictions made by TRP in an ecologically valid setting. In Shanghai, it is common for people to be offered jobs on a fixed salary, a flexible (high variance) salary, or on an intermediate (low variance) salary. The authors asked college seniors how much they would need to earn in order to survive in Shanghai (MR) as well as what they hoped to earn in their first job (G). In accordance with our findings, Wang and Johnson showed that college students chose payment plans depending on their ability to surpass these real-world reference points.

Even more aptly, our results show that behavioral data from animal foraging represents constructs similar to those found in human beings (Hurly, 2003). Hurly showed that hummingbirds became more risk averse when approaching a survival requirement from above (MR) and risk seeking when approaching a reproductive threshold from below (G). Although our experiments may not have high ecological validity, other research strongly supports the applicability of our results to everyday settings.

Future directions

Despite this support, more research should be done to bring research regarding multiple reference points into reasonably lifelike paradigms. Specifically, these paradigms could include work with setting household budgets or with goals for student grades during the semester. Using similar paradigms, it will not be possible to completely avoid imposing reference points; however, working under mandated quotas is commonplace. Overall, these reference points should be made less artificial than they were in the present experiments.

When reference points or quotas are imposed, individuals will often have autonomy in deciding either where to set these reference points or in what manner they will strive to achieve them. Research into aspirations has shown performance improves when participants are asked to set their own specific, challenging goal (Heath et al., 1999). Future research giving participants the opportunity to set their own goals will not only increase their interest in the experiment thereby providing better data, but could also show an increase in the use of multiple reference points. By providing more autonomy in realistic scenarios, it should be possible to increase both the effect size and descriptive capabilities of mRP theories.

It could also be the case that the current study has not exhausted the possible reference points that may impact choice behavior or risk preferences. For example, beyond the rigid structure imposed by distinct reference points such as G, MR, and SQ, one's expectations may also exert an influence and serve as a natural referent for comparison. Consider a situation where one has a sales target (G) and quota (MR) for a given month. If one is already above the quota, and feels that the target is realistically out of reach, then he/she may set an implicit or explicit expectation that colors the evaluation of outcomes. The subjective nature of this reference point suggests it may even relate to a "personal goal" such as mentioned in the context above.

Finally, although in Experiment 1 the mRP-predicted choice option had lower expected value, the difference in EV (ca. 50 lira) was held relatively constant across trials. As expected, the increase in subjective value achieved in crossing a reference point overcame differences in EV. However, by manipulating the difference in EV across gambles in future studies, we could begin to assess the limits of EV differences that will still result in choosing a gamble that enables surpassing a reference point.

DOI: 10.1002/bdm

Our findings demonstrate that multiple reference points (a minimum requirement, status quo, and goal) can all impact behavior within the same choice context. Rather than directly competing with prospect theory, our findings suggest the need to integrate many of the basic principles initially put forth by Kahneman and Tversky (1979) as well as the critiques of prospect theory contained in mRP-dependent accounts (e.g., SP/A theory, Lopes & Oden, 1999; TRP theory, Wang & Johnson, 2009. We are confident that future research incorporating more lifelike contexts will continue to affirm the increase in descriptive power of multiple-reference-point dependence as a model for real-world decision scenarios.

APPENDIX A: MANIPULATION CHECK FOR STRONG REFRENCE POINT CONDITION IN EXPERIMENT 1, AND EXPERIMENT 2

How many lira must you win today in order to receive *any* entries into the bonus drawing?_____
How many lira must you win today in order for your number of entries to be doubled?_____
If you win 800 lira today, how many entries would you earn in the bonus drawing?_____
If you win 2800 lira today, how many entries would you earn in the bonus drawing?_____
If you win 4800 lira today, how many entries would you earn in the bonus drawing?_____

APPENDIX B: MANIPULATION CHECK FOR WEAK REFERENCE POINT CONDITION IN EXPERIMENT 1[10]

Please enter your answers to the following questions in the spaces provided.

What is the range of possible values for the lower goal?

____________ To ____________

What is the range of possible values for the upper goal?

____________ To ____________

The most likely value for the lower goal is… ____________

The most likely value for the upper goal is… ____________

If you win 800 lira today, how many entries would you earn in the bonus drawing if…

… the lower goal is randomly selected to be 700? ____________

… the lower goal is randomly selected to be 900? ____________

If you win 4,200 lira today, how many entries would you earn in the bonus drawing if…

… the upper goal is randomly selected to be 4,100? ____________

… the upper goal is randomly selected to be 4,300? ____________

[10]Participants were also provided with a second page providing the image included in this manuscript as Figure 1.

DOI: 10.1002/bdm

REFERENCES

Bell, D. (1982). Regret in decision making under uncertainty. *Operations Research*, *30*, 961–981.

Borch, K. (1968). Decision rules depending on probability of ruin. *Oxford Economic Papers*, *20*, 1–10.

Bordley, R., & LiCalzi, M. (2000). Decision analysis using targets instead of utility functions. *Decisions in Economics and Finance*, *23*, 53–74.

Brown, D. R. (1953). Stimulus similarity and the anchoring of subjective scales. *American Journal of Psychology*, *66*, 199–214.

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

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, *38*, 79–109.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Hurly, A. T. (2003). The twin threshold model: Risk-intermediate foraging by rufous hummingbirds, Selasphorus rufus. *Animal Behaviour*, *66*, 751–761.

Jeffrey, S. A., Onay, S., & Larrick, R. P. (2010). Goal attainment as a resource: The cushion effect in risky choice above a goal. *Journal of Behavioral Decision Making*. *23*, 191–202.

Kacelnik, A., & Bateson, M. (1997). Risk-sensitivity: Crossroads for theories of decision making. *Trends in Cognitive Science*, *1*, 304–309.

Kahneman, D., & Miller, D. T. (1986). Norm theory – Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

Kahneman, D., & Tversky, A. (1979). Prospect theory. *Econometrica*, *47*, 263–292.

Loomes, G., & Sugden, R. (1982). Regret theory—An alternative theory of rational choice under uncertainty. *Economic Journal*, *92*, 805–824.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology*, *20*, 255–295.

Lopes, L. L., & Oden, G. C. (1999). The role of aspiration level in risky choice: A comparison of cumulative prospect theory and SP/A theory. *Journal of Mathematical Psychology*, *43*, 286–313.

March, J. G. (1988). Variable risk preferences and adaptive aspirations. *Journal of Economic Behavior and Organization*, *9*, 5–24.

March, J. G., & Shapira, Z. (1992). Variable risk preferences and the focus of attention. *Psychological Review*, *99*, 172–183.

Markowitz, H. (1952). The utility of wealth. *The Journal of Political Economy*, *60*, 151–158.

Mellers, B.A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory. *Psychological Science*, *8*, 423–429.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.

Neale, M. A., & Bazerman, M. H. (1991). *Cognition and rationality in negotiation*. New York: Free Press.

Oden, G. C., & Lopes, L. L. (1997). Risky choice with fuzzy criteria. In R. W. Scholz , & A. C. Zimmer (Eds.), *Qualitative aspects of decision making* (pp. 56–82). Lengerich, Germany: Pabst Science Publishers.

Olson, J. M., Roese, J. J., & Zanna, M. P. (1996). Expectancies. In Higgins , E. T. & Kruglanski A. (Eds.), *Social psychology: Handbook of basic principles* (pp. 211–238). New York: Guilford Press.

Ordóñez, L. D. (1998). The effect of correlation between price and quality on consumer choice. *Organizational Behavior and Human Decision Processes*, *75*, 258–273.

Ordóñez, L. D., Connolly, T., & Coughlin, R. (2000). Multiple reference points in satisfaction and fairness assessment. *Journal of Behavioral Decision Making*, *13*, 329–344.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica*, *20*, 431–449.

Sanbonmatsu, D. M., Kardes, F. R., & Gibson, B. D. (1991). The role of attribute knowledge and overall evaluations in comparative judgment. *Organizational Behavior and Human Decision Processes*, *48*, 131–146.

Sullivan, K., & Kida, T. (1995). The effect of multiple reference points and prior gains and losses on managers' risky decision making. *Organizational Behavior and Human Decision Processes*, *64*, 76–83.

Tryon, W. W. (1994). Expectation. In Ramachandran V. S. (Ed.), *Encyclopedia of Human Behavior Vol. 2* (pp. 313–319). San Diego, CA: Academic Press.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Wang, X. T. (2002). Risk as reproductive variance. *Evolution and Human Behavior*, *23*, 35–57.

Wang, X. T. (2008). Risk communication and risky choice in context: Ambiguity and ambivalence hypothesis. *Annals of the New York Academy of Sciences*, *1128*, 78–89.

Wang, X. T., & Johnson, J. G. (2009). *Tri-reference point theory of risky choice*. Manuscript submitted for publication.

DOI: 10.1002/bdm

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, *12*, 93–106.

Authors' biographies:

Gregory J. Koop is a graduate assistant in the department of psychology at Miami University. His research interests include embodied decision making and computational models of decision making.

Dr. Joseph G. Johnson is an assistant professor in the department of psychology at Miami University. His research interests include computational models of decision making, embodied decision making, and process measurement.

Authors' addresses:

Department of Psychology, Miami University, Oxford, OH 45056.

Department of Psychology, Miami University, Oxford, OH 45056.

DOI: 10.1002/bdm