UNREALISTICALLY OPTIMISTIC CONSUMERS:
A SELECTIVE HYPOTHESIS TESTING ACCOUNT FOR OPTIMISM IN
PREDICTIONS OF FUTURE BEHAVIOR

by

Robin J. Tanner

Department of Marketing, Fuqua School of Business
Duke University

Date:_______________________

Approved:

___________________________

Kurt Carlson, Supervisor

___________________________

Jim Bettman, Supervisor

___________________________

Larry Sanna

___________________________

Rick Larrick

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Business Administration at the Fuqua School of Business
of Duke University

2008
ABSTRACT

UNREALISTICALLY OPTIMISTIC CONSUMERS:
A SELECTIVE HYPOTHESIS TESTING ACCOUNT FOR OPTIMISM IN
PREDICTIONS OF FUTURE BEHAVIOR

by

Robin J. Tanner

Department of Marketing, Fuqua School of Business
Duke University

Date:_______________________

Approved:

___________________________
Kurt Carlson, Supervisor

___________________________
Jim Bettman, Supervisor

___________________________
Larry Sanna

___________________________
Rick Larrick

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Business Administration at the Fuqua School of Business of Duke University

2008
ABSTRACT

Individuals tend to make unrealistically optimistic self assessments about themselves and their future behavior. While little studied in marketing, unrealistic optimism by consumers may have negative consequences for both marketers and consumers. This dissertation proposes and explores a selective hypothesis testing view of unrealistic optimism. Specifically, I propose that consumers adopt the tentative hypothesis that they will behave in an ideal fashion when predicting their future behavior. They then selectively test this hypothesis by accessing information consistent with it, with the ultimate consequence being unrealistically optimistic predictions of future behavior.

To validate this theory I use the following experimental paradigm. I have individuals first provide an idealized estimate for the behavior of interest (e.g., In an ideal world, how often would you exercise next week?) and then provide a second estimate (e.g., How often will you exercise next week?). The idea here is that by making the idealized nature of the ideal behavior salient consumers will be less likely to test a hypothesis of ideal behavior when subsequently providing an estimate. In a series of ten studies, I find that prior consideration of idealistic performance does indeed temper optimism in subsequent self-assessments (henceforth post-ideal estimates). Specifically, post-ideal estimates are free of relative optimism versus expectations of others.
behaviors, are more reflective of actual past behavior, and better predict actual future behavior. Furthermore, this attenuation of optimism is mediated by increased consideration of realistic thoughts and is moderated by both expertise and decisiveness. All of these results are consistent with selective hypothesis testing being a key driver of unrealistic optimism. Additionally I demonstrate that the debiasing effect of my method extends from behaviors to above average (and in some cases below average) views of traits and abilities. As such my work raises the possibility that selective hypothesis testing underlies a wide variety of self assessment biases.

Having found strong support for my selective hypothesis testing view of unrealistic optimism, I also explore the potential consequences that unrealistic optimism may have for consumer decisions. In particular, I demonstrate that unrealistically optimistic predictions of future behavior appear to be associated with greater willingness to pay for socially desirable products (e.g., treadmills) and that attenuation of such optimism can reduce willingness to pay. Some researchers have argued that unrealistic optimism with respect to future behavior causes people to make vice choices in the present because they expect to make virtuous choices in the future (Kahn and Dhar 2007). If so, then the current research suggests one way to help consumers from falling into the trap of justifying vice behaviors with optimistically held views about future actions.
TABLE OF CONTENTS

Abstract......................................................................................................................... iv

List of Tables.................................................................................................................. viii

List of Figures................................................................................................................ ix

Introduction.................................................................................................................... 1

Background..................................................................................................................... 3

Unrealistic Optimism....................................................................................................... 3

Consequences of Unrealistic Optimism........................................................................... 5

A Selective Hypothesis Testing Account of Unrealistic Optimism................................. 8

Experimental Overview............................................................................................... 13

Study 1: Initial demonstrations of attenuation of optimism........................................ 14

Study 2: Evidence of objective accuracy....................................................................... 20

Study 3: Testing for a Dual response confound............................................................. 27

Study 4: Explicit Instruction to Make Realistic Predictions......................................... 29

Study 5: Alternative Wording....................................................................................... 32

Study 6: Moderation by Decisiveness........................................................................... 34

Study 7: Willingness to Pay for a Product................................................................. 36

Study 8: Product Utilization and Willingness to Pay................................................... 40

Study 9: The Above Average Effect............................................................................ 46

Study 10: The Below Average Effect.......................................................................... 48

General Discussion....................................................................................................... 50

Contributions to Theory.............................................................................................. 51

Implications for Consumer Choice and Welfare......................................................... 54
LIST OF TABLES

Table 1: Behavioral Estimates by Condition............................................. 18
Table 2: Estimated Task Completion Times............................................. 22
Table 3: Estimated Ability Ratings......................................................... 49
LIST OF FIGURES

Figure 1: Selective Hypothesis Testing Model..................................................12

Figure 2: Mediation by Procrastination Related Thoughts..............................26

Figure 3: IPOD WTP by Expertise Condition..................................................44
INTRODUCTION

Consider a middle-aged man who plans to start working out more frequently next month. In the face of this expectation, he might allow himself a second dessert or he might buy a home exercise bike. However, if his exercise expectations are optimistically biased, then the second dessert may have been ill consumed and the money spent on the exercise bike largely wasted. If so, then the case can be made that this consumer may have made better decisions if he could have more realistically assessed his future behavior, acknowledging perhaps that he lacked the discipline to execute a sustained home work out program. In light of this realization, he might have spent his money more intelligently.

This research focuses on consumers’ predictions of their future behavior, which I find are both unrealistically optimistic and influential on downstream judgments and decisions. More importantly, I trace the origin of these unrealistically optimistic expectations to a process of selective hypothesis testing, wherein consumers initially adopt a tentative hypothesis of idealistic future behavior and then selectively recruit information to support it. From this view, I predict that estimates of a specific future behavior should be nearly as optimistic as estimates of the same behavior made assuming ideal conditions. Moreover, it should be possible to get consumers to abandon their tentative hypothesis by making salient the contrast between ideal and real. That is,
consumers who first estimate their behavior in an ideal world (i.e., test a hypothesis of ideal behavior) should realize for themselves that the world is not ideal and so should test a more realistic hypothesis when estimating their actual behavior. Consequently, these individuals should provide more realistic estimates of actual behavior than individuals who simply estimate their behavior (i.e., without initially estimating it in an ideal world).

The remainder of this document is structured as follows. First, I provide an overview of research on unrealistic optimism and discuss the potential consequences that such optimism may give rise to. I then describe the empirical and theoretical underpinnings of a selective hypothesis testing model of unrealistic optimism. Next, I lay out a technique designed both to validate the proposed selective hypothesis testing account and to help consumers arrive at more realistic estimates of future behavior. I then present data from ten studies in which this technique is applied to the prediction of a wide range of behaviors. Finally, I conclude with discussions of the theoretical advances made by the work, its consequences for consumer choice and welfare, and with ways in which it might be extended in future research.
Unrealistic Optimism

A wealth of research has found that people view the future through rose-tinted glasses, believing that they are more likely than others to undertake positive behaviors or to experience positive events and are less likely to experience negative events or to undertake negative behaviors (Armor and Taylor 2002; Burger and Burns 1988; McKenna 1993; Perlof and Fetzer 1986; Taylor and Brown 1988; Weinstein 1980). Such unrealistic optimism effects (sometimes called comparative optimism or illusory optimism effects) occur whether people are making explicitly comparative judgments (e.g., “Compared to an average student, how likely are you to catch an STD in the next year?”) or absolute judgments (e.g., “How likely are you to catch an STD in the next year?”). They exist in situations where individuals have substantial influence over outcomes (behavioral predictions) as well as in settings where fate plays a more significant role (likelihood estimates). Examples of the behavioral predictions include optimistic estimates that one will get a good first job (Weinstein 1980), avoid an unwanted pregnancy (Burger and Burns 1988), and avoid addiction to drugs (Weinstein 1987). Examples of optimistic likelihood estimates include unrealistically low estimates
for the likelihood one will contract lung cancer (Perlof & Fetzer 1986), be in a car accident (McKenna 1993) or be the victim of a crime (Perlof & Fetzer 1986).

The robustness of optimistic bias has spawned numerous explanations, which can be broadly categorized as either motivational or nonmotivational. Motivational accounts presume that optimistic bias supports a self-serving motivation, such as the desire to avoid the anxiety associated with an unpleasant future (Hoorens 1995; Taylor and Armor 1996; Weinstein 1980). For example, an individual undertaking occasional high-risk sex might be motivated to reduce anxiety about contracting an STD, with biased beliefs about his or her relative vulnerability being invoked in the service of this motivation. Similarly, Wills (1981) contends that individuals selectively compare themselves with targets that produce esteem enhancing estimates (e.g., a comparison to individual at higher risk of contracting at STD) to support the desired conclusion that one is personally not particularly at risk.

In contrast, nonmotivated accounts contend that cognitive processes are sufficient to produce optimistically biased estimates. In this vein, Weinstein (1987) argues for an egocentrism account and suggests that individuals arrive at unrealistically optimistic self-estimates because they focus on the causal actions they will take in future scenarios, while failing to recognize similar actions that others might take. Similarly, Kahneman and Tversky (1979) argue for a self centric mechanism when they suggest that the planning fallacy is caused by individuals taking an overtly singular or “inside”
perspective, focusing on positive behaviors, while failing to consider less optimal behaviors. They conclude (Kahneman and Tversky 1982) that mental simulations about future scenarios appear to be biased toward the positive, with an “idealized” version of the future often emerging.

Consequences of Unrealistic Optimism

With unrealistic optimism being such a widespread and robust phenomenon, considerable research has been devoted to assessing the consequences of such illusory beliefs. A significant body of research suggests that unrealistic optimism can have positive consequences for mental health and/or individual welfare. It has been suggested that optimism can increase motivation, raise aspirations and strengthen coping mechanisms in the face of adversity (Taylor and Brown 1988). In particular, motivational theories suggest that unrealistic optimism motivates higher outcomes than would otherwise be achieved in the absence of the optimistic bias (Taylor and Brown 1988). For example, Scheier et al. (1989) observed that optimistic heart-by-pass patients made substantially superior recoveries compared to their less optimistic peers. However, other researchers have argued that optimistic expectancies and wishful thinking can have negative consequences. For example, unrealistic optimism has been shown to lead to the pursuit of unreasonable goals (Kahneman and Lovallo 1993),
distraction from the formation of implementation plans (Oettingen 1996), flawed judgment and disengagement (Robins and Beer 2001), poor decision making (Baumeister 1989; Becker 1974; Wood 1989), and suboptimal negotiations (Neale and Bazerman 1985). Indeed, one only needs to consider the much documented consequences of the planning fallacy (Buehler, Griffin, and Ross 1994) to find a salient example of how specific optimistic expectations (e.g. in delivery and/or project completion dates) can have negative consequences for partners, suppliers, and customers who rely on the optimistically provided estimates.

As far as explicitly consumption oriented behaviors are concerned, there is mounting evidence that consumers’ current preferences are influenced by the choices they have made in the past (Dhar, Huber, and Khan 2005) and the choices they expect to make in the future (Kahn and Dhar 2007). Moreover, it appears that consumers prefer to balance choices over time to service multiple goals (Dhar and Simonson 1999; Read, Loewenstein, and Rabin 1999). For example, a consumer may allow himself a tasty but unhealthy lunch because he plans to eat a healthy dinner. Such balancing may maximize welfare by providing a tradeoff between vices and virtues (Wertenbroch 1998) or between hedonic and utilitarian goals (Dhar and Wertenbroch 2000) that leave the consumer both physically well and hedonically gratified.

Once the hedonic goal has been satisfied, the key to achieving balance lies in following through on the utilitarian goal. Unfortunately, consumers seem to make a
practice of failing to live up their optimistically held intentions. For example, there is evidence that optimistic bias is at the root of high-risk sexual activity (Sheer and Cline 1994) and failures to immunize (Larwood 1978). In consumer choice, Kahn and Dhar (2007) have made a similar argument. They claim that when making a choice, consumers often do so expecting to make related choices in the future, and that many self-control failures result from unrealistic optimism regarding future behaviors. That is, consumers overestimate the likelihood with which they will engage in virtuous behaviors in the future, which helps them justify nonoptimal vice behaviors in the present. Then when they fail to follow through on these optimistic predictions, they are left with a sequence of vice behaviors that erode their welfare. For example, recall our middle-aged man who has eaten too many desserts that he never worked off and owns too many pieces of workout equipment that he never uses. If he had been able to realistically assess his future exercise behavior, acknowledging that he lacked the discipline to execute a home work out program, then he might have spent his money more intelligently, perhaps on fruit salad, a membership at a gym that has fitness trainers on staff, or on both.

In sum, while unrealistic optimism at times may benefit individuals from a motivational and mental health point of view, it may also have informational consequences that negatively impact decision processes in the present. This may especially be the case when individuals rely on specific estimates of future behaviors when making behavioral choices in the current period. As such I argue that it is
important to find ways to help consumers more realistically assess their future behaviors and actions.

A SELECTIVE HYPOTHESIS TESTING ACCOUNT OF UNREALISTIC OPTIMISM

The estimation of one’s future behavior involves the assembly and integration of various elements of information about one’s past behavior, one’s expected future state of mind and body, and the likely state of the world. Theories of how this assembly and integration occur vary from data-driven (i.e., bottom-up theories) to hypothesis-driven (i.e., top-down theories). In reality, the process is almost certainly a mixture of the two, with salient data giving rise to a working hypothesis that guides subsequent information acquisition and integration. For example, when estimating one’s future behavior, the working hypothesis might take the form of a desirable answer to the question about one’s future behavior. The top-down influence of such a working hypothesis greatly simplifies the task of producing an estimate. That is, by focusing on the working hypothesis and those similar to it, the judge can greatly reduce the set of information that is perceived as relevant for the estimation at hand. Individuals who simplify processing in this way are said to be engaging in selective hypothesis testing (for a review see Sanbonmatsu et al. 1998).
In line with the above, it has long been argued that consumer beliefs are often constructed around tentative hypotheses (Hoch and Deighton 1989). Thus, consumers faced with predicting their future behavior are likely to do so by constructing estimates around tentative hypotheses of their future behavior. Because there are many possibilities for one’s future behavior, consumers often simplify matters and focus on a single tentative hypothesis. I propose that the desire to think positively about oneself, which dominates motivational theories of unrealistic optimism, causes consumers to adopt a hypothesis of ideal behavior. That is, consumers adopt the hypothesis that they will behave as they would if there were no substantial obstacles preventing them from behaving as they wish to. For example, if asked to predict whether he will attend a forthcoming blood drive, an individual might adopt a tentative hypothesis that he will indeed have the time and motivation needed to donate blood.

Recent research supports this contention. Newby-Clark (2005) had participants provide best, worst, and realistic scenarios as to how they would change their exercise behavior over the next month and found that the realistic estimates most resembled best case estimates. In a similar vein, Williams and Gilovich (2006) observed across a range of domains that people’s estimates of their typical performances were similar in magnitude to their estimates of their best performances.

Building from theories of selective hypothesis testing (Gettys and Fisher 1979; Sanbonmatsu et al. 1998), I posit that consumers selectively test their (working) ideal
hypothesis by focusing on evidence that supports it and neglecting evidence that contradicts it. For example, in the blood donation example, individuals would gather information that supports their donating blood (such as past memories of donation, or anticipation of the philanthropic satisfaction that donation will bring) and neglect information that is discordant with the hypothesis (such as their tendency to procrastinate or thoughts about the pain of the needle). Indeed, research suggests that consumers access information from memory in a way that is consistent with selective hypothesis testing (Deighton 1984; Posavac, Sanbonmatsu, Kardes, and Fitzsimons 2004). The ultimate consequence of this proposed process is that predictions of future behavior will be unrealistically optimistic (e.g., overstatement of blood donation likelihood).

Though studied extensively since Wason’s (1971) initial demonstrations, very little is known about how to reduce selective hypothesis testing in situations where judgments are made quickly. That is, while there is an emerging consensus that selective hypothesis testing can be exacerbated by adding time pressure or increasing cognitive load (Cronley et al. 2005; Smith et al. 2007), I know of no empirical evidence indicating that selective hypothesis testing can be turned off. Since one of my primary objectives is to explore whether consumers can be led to make more realistic predictions of future behavior, I required a different approach. Specifically, I examine whether I can help individuals think more realistically about future behavior not by turning off selective
hypothesis testing, but instead by changing the hypothesis that is being tested. To this end, I developed the two-condition design laid out in figure 1 below.

Participants assigned to the standard request condition are asked to answer a focal question of interest pertaining to an absolute estimate of a future behavior (e.g., “How many times will you exercise next month?”). Before answering this focal question, participants in the ideal-first condition answer a related question that explicitly requires them to provide an idealistic estimate to the focal question (e.g., “In an ideal world, how many times would you exercise next month?”).

The ideal world question accomplishes two things. First, responses to it should be idealistic, so I can use them to determine the extent to which standard condition responses are themselves idealized. Second, and more importantly, responding to this question should make salient to individuals that their ideal behavior is dissimilar to their likely actual behavior. This should subsequently lead them to test a more realistic hypothesis as regards their future behavior. As a consequence, participants in the ideal-first condition should selectively access thoughts that are inconsistent with ideal behavior, resulting in more realistic responses to the focal question than provided by participants in the standard question condition.
Hence, I expect to observe the following pattern of results: Ideal world estimates in the ideal-first condition should be similar to focal question responses in the standard request condition. Additionally, responses to the focal question by those in the ideal-first condition should be more realistic than those of participants in the standard request condition. Finally, with optimism largely attenuated for participants in the ideal-first condition, downstream decisions should be improved.
Experimental Overview

In what follows, I present data from ten studies designed to explore the consequence of contemplating idealistic behavior before making predictions of actual behavior. The domains in these studies span a wide range of behaviors, including blood donation, exercise frequency, willingness to pay for a product, completion time for a project, and saving for retirement. The first study (which consists of three experiments) provides initial empirical demonstrations of the basic effect of the two-question approach. The second study (containing two experiments) examines actual subsequent behavior and demonstrates that predictions made by participants in the ideal first condition are genuinely more accurate. Next I present four studies designed to examine particular aspects of the model and to rule out certain alternative explanations. The next two studies investigate whether the apparently more realistic estimates elicited by the two-question technique affect related downstream decisions, such as willingness to pay for products. Finally, I conclude with two studies which begin to explore to what extent the selective hypothesis testing model may generalize to self assessment biases beyond behavioral predictions per se (such as to above average effects).
STUDY 1: INITIAL DEMONSTRATIONS OF ATTENUATION OF OPTIMISM

This first study applies the two-question technique to estimates of future behaviors in the domains of blood donation likelihood, weekly exercise frequency and savings discipline. The studies were run separately, but they are reported together to avoid repetition in the description both of their common method and in the pattern of results.

The particular domains were chosen as beliefs about how one will behave in them are potentially consequential from a choice and welfare point of view. For example, unrealistic optimism regarding future exercise behavior has clear potential to negatively influence individuals’ choices. If individuals optimistically believe they will exercise more in the future than they actually will, then in the current period they may exercise less, consume more unhealthy foods, or overspend on items that will go underused (e.g., exercise equipment or work out attire). As far as blood donation is concerned, statistics show that roughly 25% of Americans will require blood at least once in their lifetime. However, in any given year less than 5% of Americans donate blood. The reality is that a minority of Americans are essentially providing insurance (by keeping our blood supply viable) for the rest of us. Despite this, when asked whether they will donate blood at an upcoming blood drive, substantially greater than 5% of
people say they will do so (see below). However, as the data above indicate, most people fail to follow through on their intentions.

The final domain, retirement savings, is one where optimistic expectations of future performance appear particularly rife. A 2006 Employee Benefit Research Institute report found that 69% of workers were somewhat to very confident about their prospects for financial security in retirement. Despite this, 65% of all workers said they had less than $50,000 in total savings and investments, while 22% of the very confident group (whom one might presume to have saved the most) had not even started saving at all. The apparent tenacity of unrealistic optimism in beliefs about future financial circumstances made it a particularly attractive domain to explore.

Method

Participants were 76 undergraduates (in the blood donation experiment), 176 adult attendees at a dance performance (in the exercise experiment), and 35 undergraduates (in the savings experiment). The same basic design applied in each experiment. Participants were randomly assigned to either the standard request or ideal-first condition and thus answered just the focal question or the ideal world and focal question, respectively. In each case, participants read a brief scenario introduction before proceeding to the questions of interest. The specific questions asked in each domain are
shown below. Participants in the ideal-world conditions were presented with both questions (pertaining to a hypothetical forthcoming blood drive), while standard request condition participants were faced only with the second question.

Blood donation questions:\(^1\)

1. In an ideal world would you donate blood?
2. Will you donate blood?

Exercise frequency questions:

1. In an ideal world, how many times per week (on average) would you exercise in the next month?
2. How many times per week (on average) will you exercise in the next month?

Savings discipline questions:\(^2\):

1. In an ideal world, how disciplined would you be about saving for retirement over the course of your career?
2. How disciplined will you be about saving for retirement over the course of your career?

Participants in both the blood donation and exercise studies were additionally asked to estimate the average performance of their peers on the same dimension.

Specifically, in the exercise study participants in a third peer estimate condition

\(^1\) Measured on a binary yes/no scale.
\(^2\) measured on a 1 (not very disciplined) to 11 (very disciplined) scale.
estimated how many times the typical dance performance attendee would exercise per week over the next month. In the blood donation study participants in both conditions answered an additional question requiring them to estimate what percentage of their peers would donate blood.

Results and Discussion

The results from all three studies are displayed in table 1. A consistent pattern of results was observed across all three studies. First, in each domain there was no difference between the focal question estimates in the standard request condition and the ideal-world estimates in the ideal-first condition. This result is consistent with my contention that individuals tend to selectively test a hypothesis of ideal behavior. Second, in both the exercise and blood donation studies (in which peer estimates were obtained) focal estimates in the standard request condition were reliably higher than estimates of peer performance, a result indicative of a comparative optimistic bias. Of note, prior research has found that optimistic biases tend to stem from overly charitable self views, not from denigrated impressions of one’s peers (Epley and Dunning 2000). Hence, self estimates that are higher than peer estimates are likely to be indicative of absolute unrealistic optimism (i.e., as measured against real behavior) rather than simply being comparatively optimistic versus peers. In sum, participants in the standard
request condition provided estimates of their own behavior that were both higher than peer estimates and were no different from explicitly idealized estimates provided by a separate group. These estimates thus appeared to be optimistically biased.

**TABLE 1: BEHAVIORAL ESTIMATES BY CONDITION**

<table>
<thead>
<tr>
<th></th>
<th>Blood Donation Likelihood</th>
<th>Weekly Exercise Frequency</th>
<th>Savings Discipline (1-11)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Focal Estimates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Condition</td>
<td>69.8% ae</td>
<td>4.72 cb</td>
<td>6.9 d</td>
</tr>
<tr>
<td>Ideal-first Condition</td>
<td>44.1% a</td>
<td>3.73 c</td>
<td>5.3 d</td>
</tr>
<tr>
<td><strong>Peer Estimate</strong></td>
<td>45.4% *e</td>
<td>3.33 b</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Ideal World Estimate</strong></td>
<td>69.7%</td>
<td>5.03</td>
<td>7.1</td>
</tr>
</tbody>
</table>

a, b, c Significantly different, p < .05  
d Marginal difference, p < .06  
e No overlap in 95% confidence intervals (peer estimate continuous, self estimate binary)  
f For each study the ideal world estimate did not differ from the standard condition estimate (all p > 0.4)  
* The estimated percentage of peers who would donate blood in the ideal-first condition (M = 41.4%) did not differ from that made by participants in the standard response condition (M = 48.5%; t(74) = 1.37, p > .17); hence a weighted average is reported.

Analysis of focal question estimates made by participants in the ideal-first conditions, however, revealed a different pattern of data. In all three domains focal question estimates in the ideal-first condition were reliably lower than focal estimates in the standard request condition. Furthermore, in the blood donation and exercise studies, focal estimates in the ideal-first condition did not differ from average estimates of peer
performance. In other words, participants in the ideal first conditions predicted personal performances that were both below those of standard condition participants and which did not differ from predictions of peer performance. Hence, blood donation estimates, exercise frequency estimates, and predictions of savings discipline appeared to be less optimistically biased in the ideal-first conditions.

The data above are consistent with several of the predictions that can be derived from a selective hypothesis testing perspective on unrealistic optimism. In particular, I repeatedly observed that ideal-world behavioral estimates do not differ from standard estimates. This is consistent with the idea that participants in the standard condition are selectively testing a hypothesis of ideal behavior. Additionally, answering the ideal world question before answering the focal question caused participants to give what appeared likely to be more realistic estimates (i.e., they were lower and closer to peer estimates), which is consistent with their having tested a more realistic hypothesis. However, to more unambiguously demonstrate that post ideal estimates are genuinely more realistic it would clearly be helpful to compare the predictions made to actual subsequent behavior. Study 2 was designed with this objective in mind.
STUDY 2: EVIDENCE OF OBJECTIVE ACCURACY

People routinely underestimate how long it will take them to finish tasks. This “planning fallacy” has been documented in everything from long-term construction projects to much briefer laboratory tasks (Buehler, Griffin, and Ross 1994; Kahneman and Tversky 1979; Newby-Clark et al. 2000). The tendency to underestimate how long it will take to complete a task has potential consequences both for consumers themselves and for those who rely on their estimates. Examples of the former include lost savings from failing to return rebates on time, choosing a high-interest/high-benefit credit based on the optimistic expectation that bills will be paid on time, and the cost associated with adjusting one’s activities to address overcommitment. With respect to consequences for others, there are numerous examples where the failure to complete a particular task on schedule causes delays in related projects that can result in cascading inefficiencies in resource allocation.

This second study consists of two related experiments. The primary goal of the first experiment (henceforth described as the “planning accuracy” experiment) was to ascertain if answering the ideal world question would cause participants to render longer and more accurate estimates of task completion. Assuming this is the case, a second goal was to investigate if predicting longer completion times undermines motivation to complete the task. That is, would realistic estimates give rise to longer
actual completion times or would such estimates have no effect on task completion. The goal of the second experiment (henceforth described as the “planning thoughts” experiment) was to examine whether participants in the ideal-first condition thought more realistically when answering the focal question. To examine this question, I collected retrospective thought listings. Participants in both experiments estimated when they would complete a task (watching a DVD for a class project or completing a hypothetical leisure time report). In both cases I expected participants in the ideal-first condition would provide longer completion time estimates when answering the focal question.

Planning Accuracy Experiment

Method. Participants in this experiment were 95 executive MBA students who participated as part of an in-class exercise. They were given a DVD that contained a 1-hour focus group discussion, which they had to watch as a first step to completing a group project that was due in two weeks. In the same class in which the DVDs were distributed, participants predicted when they would watch the DVD. Half of the participants did so after estimating when they would ideally watch the DVD, “In an ideal world, when would you finish watching the DVD?” The other half simply estimated when they would watch the DVD. Actual task completion times were
collected two weeks later, with all participants answering the following question:

“When did you watch the DVD?”. Participants reported dates that were translated into the number of days from the day on which the DVD was distributed. Thus, the minimum possible value was 0 (watched it on the same day it was received) and the maximum value was less than 14 (watched it the day the assignment was due).

Results. Average completion time estimates and average actual completion times are presented in the left data column of table 2.

<table>
<thead>
<tr>
<th>Completion Time Estimates</th>
<th>Planning Accuracy Study (DVD)</th>
<th>Planning Thoughts Study (report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Condition</td>
<td>3.04 days&lt;sup&gt;ad&lt;/sup&gt;</td>
<td>1.5 days&lt;sup&gt;bc&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ideal-First Condition</td>
<td>4.43 days&lt;sup&gt;ae&lt;/sup&gt;</td>
<td>2.5 days&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Actual Completion Time</td>
<td>4.45 days&lt;sup&gt;e&lt;/sup&gt;</td>
<td>n/a</td>
</tr>
<tr>
<td>Ideal World Estimate</td>
<td>2.74 days&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.0 day&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

As in prior studies, there was no difference between standard and ideal world estimates, but focal question estimates by those in the ideal-first condition were significantly longer than focal question estimates in the standard condition. Of greater
importance, the average completion time estimates for those in the ideal-first condition ($M = 4.43$ days) did not differ from how long participants actually took to watch the DVD ($M = 4.45$ days). In other words, participants in the ideal-first condition gave estimates that were objectively accurate.

Next, I considered whether participants in the ideal-first condition (who gave longer predictions of completion time) took longer to watch the DVD than participants in the standard condition. I observed no evidence of this; the actual time to watch the DVD did not differ between the ideal-first and standard conditions ($t(93) = 1.02, p > .30$). This suggests that providing more realistic completion time estimates did not undermine actual performance of ideal-first participants.

Planning Thoughts Experiment

*Method.* Participants were 56 students who were told that the purpose of the experiment was to test materials for research in which participants would have to write a report on how they spent their leisure time for a specific day. They were informed that they would not actually have to write the report, but that they should complete all materials and think about the report writing task exactly as if they had to do so. The materials stated that the leisure report could pertain to any day from the upcoming Monday through Sunday. Instructions also stated that since memories of how we spend...
our time fade quickly, the report needed to be written within five days of the day they chose to chronicle.

Participants were randomly assigned to either the standard or ideal-first condition and made completion time predictions as follows. Those in the ideal-first condition answered both of the questions below, while those in the standard condition answered just the second question.

In an ideal world, when would you expect to complete the report?

When do you expect to complete the report?

After this, participants were asked to briefly describe the major thoughts they had as they answered the focal question.

Results. Completion time estimates in this experiment were derived by subtracting the date they predicted they would write the report from the earliest possible day they could write the report. Average completion times appear in the right data column of table 2. These data reveal the same pattern as we’ve seen routinely in our studies (i.e., no difference between standard and ideal world estimates, and post-ideal estimates that are more realistic than standard condition estimates). Next, I examined the thought listing data.

I reasoned that the primary reasons for completing a class project more slowly than expected would pertain to procrastination and weak motivation (i.e., putting it off until the last minute). Thus, if participants in the ideal-first condition were thinking
more realistically, thoughts relating to procrastination and/or weak motivation should be more accessible. Hence, thought listings were coded by two independent coders for mention of either procrastination or weak motivation as reasons contributing to the report completion estimate. Agreement between the raters was high ($r = .88$) and disagreements were resolved through discussion.

Analysis revealed that a greater percentage of participants in the ideal-first condition mentioned procrastination or weak motivation as an issue involved with project completion ($M = 33\%$) than did individuals in the standard request condition ($M = 7\%, \chi^2 = 5.54, p < .03$). That is, those in the ideal-first condition engaged in more realistic thinking when formulating their completion time estimates than participants in the standard condition. This result is interesting in its own right since a significant body of research suggests that individuals are loath to admit volitional responsibility for negative outcomes. For example, research on post decision counterfactuals suggests that individuals would rather construct external reasons for poor performance than assign any personal liability (Gilbert et al. 2004).

To explore the role of these thoughts more fully, I performed a mediation analysis following Baron and Kenny (1986). Figure 2 displays the regression coefficients for the key relationships in this analysis. There was a significant relationship between question condition and reported procrastination thoughts (a), question condition and
participants’ estimated report completion time (b), and procrastination thoughts and estimated report completion time (c).

FIGURE 2: MEDIATION BY PROCRASTINATION RELATED THOUGHTS

The relationship between question condition and estimated report completion time was not significant when controlling for procrastination thoughts (d), and the reduction in the beta versus the model without the mediator is significant by a Sobel test ($z = 1.95, p = .05$). Thus, there is evidence that the effect of condition on completion time estimates was mediated by realistic thoughts.
Discussion

The data from these two experiments suggest that initially estimating ideal task completion time results in subsequent completion estimates that are more accurate and that the process of generating those improved estimates is governed by more realistic thinking, not just an unthinking adjustment process. Put differently, both the estimates themselves, and the thought listings, strongly suggested that answering the ideal world question caused participants to think more realistically when answering the focal question.

**STUDY 3: TESTING FOR A DUAL RESPONSE CONFOUND**

The possibility exists that the apparent attenuation of optimism in the ideal-first conditions in the studies to date was due partly to my asking these participants to consider the topics of interest twice (i.e., to answer the ideal-world question and then the standard question). For example, perhaps when asked to consider the topic of blood donation for a second time in the ideal-first condition, participants simply thought more deeply. To directly rule out this possibility, study 3 examined whether a simple reconsideration of the focal blood donation question would itself lead to reduced donation estimates.
Method

Participants were 40 undergraduates who were paid for their participation. Each participant was randomly assigned to the standard request, ideal-first, or to a dual response control condition. Those in the ideal-first condition answered the two-questions below, while those in the standard request condition answered only the second question. Responses to both questions were captured on a scale from 0-100%.

In an ideal world how likely would you be to give blood next week?

How likely are you to give blood next week?

Participants in the dual response control condition answered the second question above and then read the following before answering this same question again

Previous research has shown that people are not always accurate when first asked for an estimate such as the above. Please take a few moments to consider again your blood donation likelihood.

Results and Discussion

Ideal world donation likelihood estimates for those in the ideal-first condition ($M = 76.1\%$) did not differ from focal question estimates of participants in the standard request condition ($M = 71.2\%; t(25) = -.38, p > .71$), suggesting again that focal question
responses in the standard condition were idealized. Also consistent with study 1, focal question estimates were significantly lower in the ideal-first condition ($M = 40.0\%$) than in the standard request condition ($t(25) = 2.20, p < .04$). Of greater import, there was no reduction in estimated donation likelihood for participants in the dual response condition. That is, responses to the second request ($M = 70.0\%$) did not differ from the first request ($M = 72.7\%, t(12) = .90, p > .39$). These data thus rule out the possibility that the observed attenuation of optimism occurred simply because participants in the ideal-first condition answered two questions about donating blood.

**STUDY 4: EXPLICIT INSTRUCTION TO MAKE REALISTIC PREDICTIONS**

Study 4 was designed to address two distinct issues. First, from a practical perspective, the two-question technique has clear potential for helping consumers make more realistic predictions of their future behavior. However, if realistic predictions could be obtained simply by instructing individuals to avoid making idealistic assessments, then the two-question technique would have less utility as a corrective technique.

Second, one alternative explanation for the results so far might be that they are a function of experimental demand. That is to say, participants might have inferred that they were supposed to give a lower estimate to the second question and were happy to oblige the experimenter in doing so. To address both of these issues, study 4 examines
whether a simple instruction to avoid idealistic predictions will by itself reduce optimistic bias. The failure of direct instruction to render a debiasing effect would clearly cast the two-question technique in a more favorable light regarding debiasing effectiveness. Furthermore, if experimental demand does underlie the results to date, an explicit request to provide a lower estimate would presumably result in similarly accommodating responses from participants.

Method

Participants were 133 graduate students who completed this and other unrelated studies in return for a charitable contribution to a graduate student club. As in study 1, the focal question in this study required participants to estimate their weekly exercise frequency over the next month. Each participant was randomly assigned to one of three conditions. The first two conditions were identical to the standard and ideal-first conditions in study 1. The explicit instruction condition was a modified version of the standard condition in which participants were additionally instructed to provide realistic estimates as follows:

Please do not provide an idealistic prediction, but rather the most realistic prediction of your behavior that you can.
Just like participants in the standard condition, participants in this explicit instruction condition read and answered only the focal question.

Results and Discussion

The data pattern for the standard and ideal-first conditions replicated the pattern from study 1, with focal exercise estimates directionally higher in the standard condition ($M = 3.35$) than in the ideal-first condition ($M = 2.39; t(64) = -1.39, p = .17$). However, more important is that focal exercise estimates in the explicit instruction condition ($M = 4.37$) were significantly higher than those in the ideal-first condition ($t(66) = 2.34, p < .05$) and were directionally higher than standard condition estimates. Not only did a direct instruction to provide a realistic estimate fail to reduce optimistic bias, it may have increased it slightly, a result which weakens potential demand accounts. This result is consistent with Self-Determination Theory (Deci and Ryan 1985), which suggests that autonomously originating motivations lead to superior performance. That is, while individuals in the explicit instruction condition were instructed to avoid idealistic estimates, those in the ideal-first condition presumably arrived at this conclusion autonomously. Since Self-Determination Theory also suggests that autonomously generated motivations lead to more persistent behavioral change, this bodes well for the potential of our approach to precipitate more than a fleeting change in self view.
STUDY 5: ALTERNATIVE WORDING

Study 5 was primarily designed to determine whether an alternative wording of the ideal world question would produce similar results as those above. This is important, because if my selective hypothesis testing account is correct it should be possible to get individuals to realize that they should not default to an idealized behavioral scenario without using the specific phrase “in an ideal world”. Additionally, this study provided an opportunity to further address the question of whether post ideal estimates are more reflective of actual behavior, in this case past actual behavior.

Method

Participants were 56 graduate students who completed this study as part of a series of class exercises. All participants estimated how many times they would exercise over the next two weeks. Two weeks later, participants reported their actual exercise frequency for the two week period. The day after reporting actual exercise frequency, participants estimated how often they would exercise in the next two week period. Half of the participants did so after answering a variant of the ideal world question designed to remove all possible constraints that might impede one’s ability to exercise (e.g., constraints on time, motivation, and physical ability). The question was worded as
follows, “If there were no constraints on your time, motivation, and physical ability, how many times would you exercise over the next two weeks?”

Results and Discussion

Because the phrasing of this question was designed to make participants consider their behavior if there was absolutely no constraints, I expected participants would answer it with extremely high estimates, which they did ($M = 10.26$). Nevertheless, I expected that answering this question would have the same general effect as the ideal world question has had in the prior studies, namely causing respondents to realize that the world has constraints, which would lead them to test a more realistic hypothesis when subsequently answering the focal question.

Participants in the standard condition estimated that they would exercise 4.93 times over the next two weeks, an estimate that was significantly greater than their actual exercise frequency over the previous two weeks ($t(82) = 2.17, p < .05$). Thus, it seems that these participants did not use recent exercise failures to update their exercise estimates. In contrast, participants in the no constraint condition estimated they would exercise just 3.70 times, an estimate that was not different from actual exercise reports for the preceding two weeks ($t(80) = 0.47, p > .60$).
From these data, I draw two conclusions. First, the effect of the ideal world question (in the experiments above) is not limited to use of the phrase “in an ideal world” because an alternative wording that emphasized the lack of constraints had a similar effect on subsequent responses to the focal question. Second, this study provides incremental evidence that post-ideal responses do not just appear more realistic, but are genuinely more reflective of actual behavior. That is, if participants’ self-reported exercise frequency (of three times per two weeks) is an accurate reflection of their typical exercise behavior, then I can conclude that getting participants to initially consider an extremely optimistic hypothesis subsequently leads to exercise estimates that are not just lower, but which are not statistically different from typical exercise behavior. In other words, this study provides further evidence that the prior consideration of ideal behavior leads to subsequent predictions of future behavior that better predict actual future behavior.

**STUDY 6: MODERATION BY DECISIVENESS**

It is worth noting that even though there were no statistically significant differences between ideal world and standard estimates in the studies above, there is a directional pattern in which standard estimates are slightly more realistic than ideal world estimates. Thus, it is possible that some individuals in the standard conditions
above tested both the ideal behavior hypothesis and a more realistic one. Consistent with this idea of heterogeneity in testing, there is evidence that consumers who are low in need for cognitive closure are more likely to test multiple hypotheses (Cronley et al. 2005; Kardes et al. 2004). Accordingly, I reasoned the decisiveness factor of Webster and Kruglanski’s (1994) need for cognitive closure scale (e.g., “I usually make important decisions quickly and confidently”) would be a good proxy for the tendency to test multiple hypotheses in behavioral predictions. That is, those low in decisiveness should be more likely to test multiple hypotheses, including some that are realistic, when answering the standard question. Hence, I expected that individuals low in decisiveness would provide standard estimates that were farther from their ideal world estimates than would their high decisiveness counterparts.

Method

Participants were 25 graduate students who completed this study as part of a series of class exercises. All participants estimated how often they would exercise over the next two weeks. Two weeks later, these same individuals estimated how many times they would exercise over the following two weeks in an ideal world. A further two weeks later, these individuals answered the seven questions that form the decisiveness factor of the need for cognitive closure scale.
Results and Discussion

I expected that the difference between ideal world estimates and standard condition estimates would increase as decisiveness decreased. That is, those who are most likely to consider multiple hypotheses when contemplating the standard question (i.e., those low in decisiveness) should give standard estimates that are farthest from their ideal world estimates. The data supported this hypothesis. Specifically, the correlation between decisiveness and the difference measure was significant and negative \( (r = -.367, p < .05) \). These data thus provide incremental support for the idea that standard condition estimates stem from a process of selective hypothesis testing. Namely, greater decisiveness (which should be associated with a greater tendency toward testing a single hypothesis) was associated with less divergence between standard and ideal world estimates.

STUDY 7: WILLINGESS TO PAY FOR A PRODUCT

Having established in Studies 1-6 that the two-question approach attenuates optimism in predictions of future behavior, I next explore whether these estimates will influence willingness to pay for products related to the behavior in question. Such a result would be important for two reasons: First, an alternative way to substantiate the
selective hypothesis testing framework is to examine downstream decisions. If ideal-first participants genuinely test a more realistic behavioral hypothesis when answering the focal question, then they consider more realistic information. As a result subsequent decisions using such information as an input should be affected. Second, I have argued that improving forecasts of future behaviors may have the potential to help improve consumer decision making. Examining the potential link between optimistic predictions and willingness to pay for products is one way to substantiate this claim.

Optimistic bias about one’s future behavior may contribute to an extremely common consumption phenomenon, namely, overbuying of products that see little actual use (Morris and Bronson 1970). Indeed, consumers who were asked why they fail to use unused products were most likely to cite reasons relating to a mismatch between their aspirations for product use and reality (Trocchia and Janda 2002). Not only does health and fitness equipment fit the stereotype for unused products, but evidence suggests that this category is indeed rife with wasted purchases (Trocchia and Janda 2002). Thus, given our focus on predictions of exercise behavior, exercise equipment is an obvious category to examine.

This study was also a replication of the exercise portion of Study 1, with one critical change. After answering the focal question, participants in both conditions were asked to report their willingness to pay for a new treadmill. If focal question estimates of exercise frequency differ across conditions because participants in the two conditions
accessed different information subsets, then willingness to pay for the treadmill should also differ. That is, participants in the ideal-first condition, who think less idealistically about their future exercise behavior, might be expected to report lower willingness to pay for a treadmill, since they would be more cognizant of the potential for it to see little use.

Method

Participants were 118 undergraduate students who were each paid $1 for completing this study. Each participant was randomly assigned to one of two conditions. Those in the ideal-first condition reported their willingness to pay for a treadmill after answering the ideal world and focal questions from study 1. Those in the standard condition did the same, after answering just the focal question.

Results and Discussion

As in study 1, exercise estimates in the standard condition ($M = 3.16$) were higher than focal question estimates in the ideal-first condition ($M = 2.47$; $t(116) = 2.20, p < .05$). Of greater interest, participants in the ideal-first condition reported a lower willingness to pay for the treadmill ($M = \$480$) than participants in the standard condition ($M = \$610$;
$t(116) = 2.00, p < .05$). This is consistent with expectations. That is, by adopting a less idealized perspective about their exercise behavior, ideal-first participants were better placed to assess their likely utilization of a new treadmill and, as a result, placed a lower valuation on it. Note that this outcome is potentially superior to one where consumers spend money on exercise equipment that goes unused. Consumers who honestly admit they will not execute a home work out program can address their goal to be fit in a manner that is more achievable for them (e.g., by eating more healthily or by joining a club that has fitness instructors on staff).

A remaining question is whether answering the ideal world question causes everyone to adjust their estimates toward the more pessimistic end of the response set. If so, then those who would have given realistic estimates without the intervention might give overly pessimistic estimates. Note, however, that the selective hypothesis testing account predicts that the ideal world question will cause only those consumers who would default to an idealistic hypothesis to adjust their behavioral estimates. In other words, estimates of those who would default to a realistic hypothesis in the standard condition should not be influenced by the ideal world question. As such, we expect that answering the ideal world question will not uniformly influence all consumers. One goal of the next study is to explore this issue.
STUDY 8: PRODUCT UTILIZATION AND WILLINGNESS TO PAY

In this study I further explore the downstream consequences of more realistic predictions that emerge from the two-question technique and investigate whether the effect of the two-question technique is moderated by domain expertise. Previous research has argued that experts are less likely to fall prey to selective processing effects than are novices (Posavac et al. 2004). Thus, I expected that the impact of the two-question technique would be reduced for experts. The main downstream measure is willingness to pay for an iPod. This is assessed after participants estimate their iPod capacity utilization (i.e., how many songs they would store on it), which is either the first question answered or is answered after a related ideal world query. I use iPod ownership as a proxy for expertise (i.e., I assume owners are more knowledgeable about their likely capacity utilization).

I have argued that individuals tend to default to selectively testing the hypothesis that they will behave similarly to an idealistic case when tasked with estimating their future behaviors. I have further argued that the success of the two-question method lies in its ability to lead participants to instead selectively test a more realistic behavioral hypothesis. If these contentions are true, then I should observe the following pattern of results: First, if experts are less likely to selectively test a hypothesis that they will behave ideally, then standard condition estimates for experts should be
lower than ideal-world estimates (while for novices I would expect them to be similar).

Second, relative to novices, experts should be less susceptible to the two-question technique and should report post ideal estimates that are closer to standard estimates. Such moderation would, if obtained, further support the selective accessibility account. Similarly, to the extent that iPod valuation is contingent on expected utilization, novices should report lower willingness to pay for an iPod in the ideal-first condition than in the standard condition, while experts should give relatively similar estimates across the two conditions.

Method

The hypothesis was tested in a two (iPod expertise: own versus do not own) by two (request condition: ideal-first versus standard) between subjects design with song storage and willingness to pay for a new iPod serving as dependent variables. Participants were 227 undergraduate and postgraduate students who were paid $1 for participating. Nine surveys were completed incorrectly, leaving 218 participants on which all analyses were performed. Each participant saw a picture of an iPod and read the text below.

20 and 60GB Models: At just over half an inch thick, the iPod fits comfortably in the palm of your hand and slips easily into your pocket – and your life. Merely
5.6 ounces, it weighs less than two compact disks, and even many cell phones.

And yet the iPod gives you a huge 60GB hard drive – big enough to hold 15,000 songs.

Participants in the ideal-first condition answered the two-questions below, while those in the standard condition answered just the second question.

In an ideal world, how many songs would be loaded on your iPod at any time?

How many songs would be loaded on your iPod at any time?

Participants in both conditions were then asked to assume they did not own an iPod and to report their willingness to pay for a 60GB model in response to the following question, “How much would you be willing to pay for this iPod?” Open-ended responses were recorded in dollars. Finally, to give a measure of expertise, participants were asked if they currently owned an iPod.

Results

For non-owners (n = 67), song storage estimates were higher for participants in the standard request condition (M = 4,854) than in the ideal-first condition (M = 1,993; t(63) = 2.76, p < .01). Furthermore, consistent with the emerging pattern from the studies to date, ideal world estimates in the ideal-first condition (M = 4,680) did not differ from standard condition estimates (p > .9). However, for iPod owners (n = 151) a different
pattern emerged. First, unlike the previous studies, ideal world estimates in the ideal-first condition ($M = 7,676$) were dramatically higher than standard condition estimates ($M = 4,184$; $t(151) = 4.00, p < .01$). Second, focal estimates in the ideal-first condition ($M = 3,469$) did not differ from standard condition estimates ($t(151) = 1.04, p = .30$). Both of these results suggest that the song storage estimates made by iPod owners were less affected by a process of selective hypothesis testing. These data are thus consistent with expectations. Namely, iPod owners have real expertise with iPod utilization, which makes them less likely to selectively test a hypothesis of ideal capacity utilization. The fact that song utilization in the ideal-first condition is higher for iPod owners than non-owners likely stems from the greater importance that music plays in the lives of iPod owners.

Next I examined the willingness to pay estimates. Figure 3 plots willingness to pay for an iPod by condition. A two-way ANOVA of willingness to pay yielded a main effect of ownership ($F(1, 214) = 22.27, p < .005$), a marginal main effect of request condition ($F(1, 120) = 3.33, p < .07$), and an ownership by request condition interaction ($F(1, 214) = 4.40, p < .04$).

Planned contrasts were conducted to expose the nature of the interaction. As expected, non-owners in the ideal-first condition reported a lower willingness to pay ($M = $129) than non-owners in the standard condition ($M = $188; $F(1, 214) = 5.42, p < .03$). However, for iPod owners, willingness to pay in the ideal-first condition ($M = $231) did
not differ from that in the standard condition \( (M = $227; p > 0.9) \). As with iPod usage, the higher willingness to pay by iPod owners probably reflects greater preference for the iPod in general. It is worth noting that this difference in average willingness to pay rules out the explanation that the null effect for iPod owners was due to a floor effect (i.e., that they were unwilling to pay anything because they already owned one).

FIGURE 3: IPOD WTP BY EXPERTISE CONDITION

![Chart showing iPod WTP by expertise condition](image)

Discussion

In this study I explored a situation where the focal question required a prediction of how fully one would use the capacity of a specific product (i.e., iPod song capacity
utilization). The results indicate that prior consideration of idealistic usage had much less effect on subsequent usage estimates provided by experts that it did for novices. This is consistent with expectations, namely that experts are relatively less likely to carry out selective hypothesis testing than are novices. Also supporting this line of reasoning was the finding that ideal-world and standard estimates were similar for novices (consistent with selective testing of a hypothesis of idealistic usage being tested in the standard condition), while for experts the ideal-world estimate was considerably higher. Thus, it seems that those who are unfamiliar with how they will interact with a product on a daily basis are more likely to report idealistic usage estimates and consequently to inflate their valuations of the product. These findings fit with the logical supposition that the consumers who are most at risk of making idealized future usage predictions, and thus to making ill advised purchases, are precisely those who lack experience with a particular product. This is both comforting and troubling. On the one hand, it is good to know that some people can separate their ideal product usage from their real usage. On the other hand, when we talk of improving consumer welfare by improving consumer choices, our silent assumption is that the consumers who need the most help are those for whom the product domain is most novel.
STUDY 9: THE ABOVE AVERAGE EFFECT

Studies 1-8 share a common theme in that they explore predictions of future behaviors. Optimism is certainly not restricted to thoughts about future behavior, however, and indeed there have been numerous demonstrations that people believe themselves to be above average in traits and abilities (Alicke 1985). It seems plausible that optimism about personal traits and abilities might be similarly caused by a default inclination to test the hypothesis that one is above average. To explore this possibility I examine perceptions of math ability. Math, which is often perceived as a difficult subject, might sometimes yield a below-average effect (Kruger 1999). However, participants in this study were undergraduate students at a university where 66% of freshman had SAT math scores over 700. Since this sample had a high absolute level of math proficiency, I expected participants would exhibit self-favoritism (i.e., the above-average effect).

Method

Participants were 78 undergraduates who were compensated for participation. Each participant was randomly assigned to either the ideal-first or standard condition. Participants in the ideal-first condition answered the following two questions:
In an ideal world, which percentile would you occupy relative to your Duke peers in math skills?

Which percentile do you occupy relative to your Duke peers in math skills?

Those in the standard condition answered only the second (i.e., focal) question. Responses were collected on a 100-point scale anchored at the bottom, middle, and top with the terms “near the bottom”, “average”, and “near the top.” I expected that the average math ability estimate by participants in the standard condition would be greater than 50, reflecting self-favoritism in the form of the above average effect. In contrast, if answering the ideal world question caused participants in the ideal-first condition to consider a more realistic subset of self-relevant information when answering the focal question, they should exhibit less self-favoritism in their estimates.

Results and Discussion

Participants in the standard condition reported math abilities that were (on average) higher than the true population average of 50 ($M = 60.1; t(38) = 2.76, p < .01$). In contrast, math ability estimates provided by participants in the ideal-first condition did not differ from the true population average ability ($M = 48.8; t(38) = .30, p > .75$), and they were reliably lower than standard condition estimates ($t(76) = 2.12, p < .05$). Thus, those in the standard condition appeared to be subject to an above average effect, while those in the ideal-first condition did not. These data provide initial support for the
effectiveness of my two-question technique in attenuating optimistic biases beyond behavioral predictions. In the next study I further explore the robustness of this technique by testing whether a variant of the approach can similarly attenuate the below average effect.

**STUDY 10: THE BELOW AVERAGE EFFECT**

Study 9 demonstrated an apparent attenuation of the above average effect. However, there are instances where self-favoritism is not the norm, but rather where people systematically report themselves as below average (Kruger 1999). Study 10 was designed to investigate whether such below-average effects might be mitigated with a variant of my approach, (i.e., one designed to cause retrieval of less pessimistic information). I explored this possibility for self reported estimates of musical ability and juggling skill.

**Method**

Three hundred and sixty undergraduate and postgraduate students were each paid $5 for participating in this and other unrelated studies. Participants were randomly assigned to either the standard or worst-first condition. Those in the worst-first
condition answered the two questions below for both juggling and musical ability, while those in the standard condition answered just the second (i.e., focal) question for each skill.

In a worst case scenario, which percentile would you occupy relative to your Duke peers in [juggling/musical] ability?

Which percentile do you occupy relative to your Duke peers in [juggling/musical] ability?

Results and Discussion

The results of this study are displayed in table 3. Average percentile ratings in the standard condition were significantly below 50% for both abilities, indicating a below average effect.

<table>
<thead>
<tr>
<th>Focal Question</th>
<th>Musical Ability</th>
<th>Juggling Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Condition</td>
<td>43.5</td>
<td>34.6</td>
</tr>
<tr>
<td>Worst-first Condition</td>
<td>50.6</td>
<td>44.9</td>
</tr>
<tr>
<td>Worst Case Question</td>
<td>44.0</td>
<td>36.8</td>
</tr>
</tbody>
</table>

Next I examined focal question responses by participants in the worst-first condition. As expected, estimates for both juggling ($M = 44.9$) and musical ability ($M = 50.6$) were significantly higher than standard condition responses (juggling: $t(236) = 2.92$,
Indeed, for musical ability, worst-first focal question estimates were not only higher than standard condition estimates but they were not reliably different from the true average of 50% \(t(121) = .25, p > .8\). Thus, the two-question approach reduced the magnitude of the below average effect for juggling and eliminated it for musical ability. Just as standard condition responses were indistinguishable from ideal world responses in earlier studies, standard condition estimates mirrored worst case responses in this study. This suggests that participants in the standard condition may have defaulted to pessimistic self assessments. More generally this symmetry of effectiveness in attenuating both above and below average effects suggests that selective hypothesis testing may in fact underlie a wide variety of self assessments. While the default hypothesis tested is normally biased in a positive direction, for a minority of cases it is overly pessimistic. The question as to what determines whether individuals test optimistic or pessimistic scenarios about their abilities and future performance certainly merits further research.

**GENERAL DISCUSSION**

Taken together, the results of ten studies demonstrate that prior consideration of idealistic behavior causes consumers to give more realistic estimates of future behavior. That is, behavioral estimates that followed prior consideration of idealistic behavior

\(p < .01; \text{musical ability: } t(236) = 1.99, p < .05\).
were consistently less idealistic than standard estimates, did not differ from estimates of peers, were more reflective of past actual behavior, and better predicted actual future behavior. I also found that participants reported having thoughts that appeared less idealistic and that they made downstream decisions that were governed by these thoughts. Finally, answers to focal questions by those who answered only the focal question were consistently similar to responses to the ideal world questions, a result that by itself suggests that people tend to answer self assessment questions from an idealistic perspective, consistent with my contention of selective testing of a hypothesis of ideal behavior.

Contributions to Theory

The current data contribute to the literature on unrealistic optimism in several ways. First, the totality of the data suggests that selective hypothesis testing is a contributor to unrealistic optimism. The evidence to support this is as follows: As noted above, standard estimates were routinely as high as ideal-world estimates. If consumers adopt a tentative hypothesis and then recruit information to test it (e.g., Deighton 1984), an explanation for the similarity between these estimates is that standard estimates are constructed around a hypothesis of ideal behavior. Moreover, if unrealistic optimism derives from selective testing of an ideal hypothesis, then it should be possible to
ameliorate it by getting individuals to move away from their default/ideal hypothesis. I believe this is precisely the effect that answering the ideal world question had on my participants. Specifically, participants who initially answered the ideal world question subsequently predicted future behaviors that did not differ from estimates of peer behavior, and were more in line with actual future behavior. Additionally, consistent with previous work (Kardes et al. 2004), I found that greater decisiveness was associated with smaller differences in standard and ideal world estimates. Finally, the expertise moderation results in study 8 are consistent with prior work that finds experts are less likely to engage in selectively hypothesis testing (Posavac et al. 2005).

Another contribution, albeit subtle, is the result that standard estimates were often indistinguishable from ideal estimates. This by itself suggests a more nuanced view of unrealistic optimism. Specifically, while it is a common finding that individuals are too optimistic in predictions of their future behavior, the current work reveals that behavioral predictions are not simply too high by some arbitrary amount, but rather often line up with perceptions of ideal behavior.

Third, I note that recent research in the areas of leader-driven primacy and selective hypothesis testing (Bond et al. 2007; Carlson, Meloy, and Russo 2006; Posavac et al. 2004) has made great strides in identifying processes whereby initial leanings or hypotheses become reinforced over time, leading ultimately to stronger preferences or views than originally existed. The current research adds to this literature on selective
hypothesis testing by extending our understanding of the types of hypotheses individuals test. That is, when estimating future behavior they appear to test the hypothesis that they will act in an idealistic fashion.

A fourth theoretical contribution of this work is that my findings lend support to biased retrieval accounts of unrealistic optimism. These accounts contrast with biased encoding views which contend that many forms of self favoritism result from the selective integration of self-favoring information into memory (Baumeister and Cairns 1992). The current data are difficult to reconcile with biased encoding because the ideal world question did not give participants information about themselves that they did not already possess.

A final contribution to the literature on unrealistic optimism comes from the planning accuracy study (see study 2). Recall that compared to participants in the standard condition, participants who predicted their completion times after answering the ideal world question gave longer and more accurate completion estimates, but they did not take any longer to actually complete the task. This suggests that the ideal world approach for getting people to adopt a more realistic hypothesis does not necessarily undermine performance. This result is noteworthy, as it contributes to the ongoing debate about the benefits and costs of unrealistic optimism (Taylor and Brown 1988), which we discuss in greater detail below.
Implications for Consumer Choice and Welfare

While the primary purpose of this research was to explore a selective accessibility explanation for unrealistic optimism, my findings may have implications for the quality of consumer decisions. In particular, is it possible that getting consumers to think more realistically can beneficially influence the quality of their decisions?

As discussed earlier research relevant to this question can be found in a number of literatures which present potentially opposing findings. In particular, a large body of research in psychology argues that optimism actually has positive consequences for mental health and/or individual welfare (Peterson 1988; Scheier et al. 1989; Segerstrom et al. 1998; Weisse 1992), while a considerable body of research in decision making literatures argues for negative consequences of unrealistic optimism (Kahneman and Lovallo 1993; Oettingen 1996; Robins and Beer 2001; Baumeister 1989; Becker 1974; Wood 1989).

At first glance, the apparent positive and negative consequences of unrealistic optimism seem to create a welfare paradox. That is, while consumers who think more realistically may make better decisions in the present, realistic thinking may also undermine their motivation to perform. In situations where motivation is undermined, the welfare consequences of more realistic thinking depend on the relative costs of lower motivation and the benefits of more realistic self assessment. While I am unqualified to
speculate on this tradeoff for any one consumer, let alone for consumers in general, I note that there may be times when the tradeoff may be unnecessary. In particular, the tradeoff is less relevant in settings where consumers can be more realistic in their predictions, without having this change in expectations undermine their motivation.

The possibility that consumers might be able to give more realistic estimates of future behavior without suffering a loss of motivation follows from a distinction between big and little optimism (Tiger 1979; Peterson 2000). Big optimism is an abstract dispositional tendency that is associated with a state of vigor and resilience. It can provide physiological and motivational benefits, such as increased desire to pursue one’s aspirations. Peterson (2000) has proposed that big optimism may underlie many of the benefits of optimism in health domains (e.g., disease recovery). In contrast, little optimism, which derives from experience with specific domains, gives rise to unrealistic optimism in those domains. As such, little optimism may contribute to suboptimal decision making in those domains (Peterson 2000).

This distinction between big and little optimism raises the possibility that helping consumers reach more realistic predictions in a specific domain need not undermine motivation to perform in that domain. Put differently, if estimates of specific future behaviors are primarily governed by little optimism, then helping consumers make more realistic predictions of future behaviors (i.e., attenuating little optimism) may improve their decisions without undermining the motivational benefits of big
optimism. The data in study 2 are consistent with this view, as the provision of more realistic estimates did not appear to damage subsequent actual performance.

Despite the evidence of study 2, it is premature to conclude that helping consumers arrive at more realistic self assessments of future behavior will never undermine performance. Simply put, it is more likely that there are some situations where being realistic undermines performance and some situations where it does not. Since the big/little optimism framework does not suggest what mechanisms might determine when performance is degraded or not, I consider the issue from a different perspective.

Prior research suggests that consumers who construct behavioral intentions sometimes treat these intentions as goals (Levav and Fitzsimons 2006; Alexander, Lynch and Wang 2008). Since goals are motivating, expectations of future behavior that are treated as goals should be motivating. If so, then a key issue in knowing whether the induction of realistic expectations will affect performance is whether consumers treat their expectations as goals. It has been suggested that abstract expectations are less goal-like (and less motivating) than concrete ones (Levav and Fitzsimons 2006). From this two possibilities emerge: First, expectations should be treated more like goals when they are concrete than when they are abstract. As such, I would speculate that expectations should be more goal-like when they are in the near future, shared publicly, given careful consideration, and linked to implementation plans. Thus, the presence of any of these
factors might cause consumers to treat their expectations more like goals, which means that causing them to adopt more realistic expectations may lead to poorer performance when these factors are present. In other words, the tradeoff between performance and realistic self appraisal is likely to be more of an issue when conditions such as these are present.

However, there is also a second possibility that stems from the idea that concrete expectations are likely to be more motivating than abstract ones. I begin by noting that ideal world expectations are (almost by definition) more abstract than real world expectations. I also note that the similarity between ideal world estimates and standard condition estimates in the studies above suggests that standard condition estimates may be the product of relatively abstract representations of the prediction context. In contrast, it is possible that individuals who answered both the ideal world question and then the standard one gave responses to the latter that were based on more concrete representations (note that my retrospective thought listing data are consistent with this conjecture). If so, then it is possible that post-ideal estimates were more goal-like than standard condition estimates. Hence, even though the former were more realistic, they may have been more motivating. For example, post ideal exercise estimates (though lower) might have been treated more like goals, and so been more likely to cause people to exercise than standard condition estimates, which though loftier, may have been treated less like goals and more like abstract ideas for the future.
In sum, it is hard to provide a definitive answer as to how increased realism might impact consumer welfare without a better understanding of precisely how behavioral expectancies influence actual behavior. As such, future research might focus on identifying the conditions which determine whether performance is degraded by more realistic estimates. I have advanced some ideas in this regard, ideas that future research should explore.

Alternative Accounts

Experimental demand (Shimp, Hyatt, and Snyder 1991) is a potential concern whenever participants might be able to guess the research hypothesis. In my design, this is mainly a concern for our finding that post-ideal focal estimates are more realistic (since the result that ideal world and standard estimates are so similar is safe from even the broadest interpretation of demand.) However, it is hard to make the case for a parsimonious demand account across all our studies.

First, it is not clear why participants would conclude that they should contrast from their ideal world estimate. Indeed, another possible deduction they could have made is that the experimenter wanted them to anchor on their ideal world estimate.

Second, even if demand does lead participants to adjust in the hypothesized direction, ideal-first participants did more than simply adjust from their ideal world
estimates. They gave focal question responses that (a) were indistinguishable both from their own peer estimates (study 1, blood donation) and from peer estimates of a control group (study 1, exercise); (b) were not different from actual recent behavior (study 1, third replicate); and (c) did not differ from actual future behavior (study 4, accuracy study).

Third, in a study where participants did not need to infer the wishes of the experimenters (i.e., where participants were told to avoid ideal estimates and to provide a realistic ones), participants did not provide estimates that were realistic. If experimental demand were at work in our studies, we should have seen very realistic estimates in this replicate.

Fourth, demand accounts do not generally extend beyond the measure of interest. However, thought listings from study 2, and the downstream willingness to pay estimates in studies 7 and 8 indicate that the effect of answering an ideal world question extended beyond the post-ideal focal question responses.

Fifth, both versions of demand are inconsistent with the expertise results in study 8. In particular, it is highly unlikely that participants could intuit on, and then deliver (in a between subjects design) the pattern of results which we observed.

Finally, I note that from a pure debiasing perspective, that is to say putting aside the psychology to which the data speak (which I strongly caution against), demand is less relevant. In other words, if helping consumers make more realistic predictions is the
sole objective, then a method’s merit should depend not on what it reveals about the process, but rather on its effectiveness.

Limitations

All of the behavioral domains I examined in the studies above arguably shared a consistent theme in that the direction individuals would be motivated to perform (and thus the likely direction in which optimism will bias predictions) is fairly clear. That is to say, we would probably all like to be able to give blood, exercise more, and complete projects speedily. While exceptions are to be expected (an Olympic athlete might desire to exercise less in an ideal world) most individuals would probably agree on which direction the average individual would be motivated to move their future performance. However, it is quite possible that in some behavioral domains this consistency of motivation across individuals might very well not be the norm. From a practical debiasing perspective such instances may have implications for the generalizability of my method.

First, it is likely that in some domains there is considerable heterogeneity in which direction individuals are motivated to change their behaviors and/or traits. For example, suppose a researcher was interested in examining predictions of their future weight made by college age males. One can imagine that overweight and skinny
members of this group are motivated to move their weight in different directions (i.e., overweight individuals down via losing fat, skinny individuals up via increasing muscle weight.) My contention is that in the presence of such mixed motivations, my method should still render a debiasing effect across all individuals. The reason for this is that since each individual has their own ideal future weight (lower for overweight, higher for skinny) then the ideal world method should work in an idiosyncratic fashion to lead each individual to revise his future weight estimate away from his ideal, be that ideal higher or lower than his current weight.

Second, and more problematic, it may be the case that in some domains optimism is less universally prevalent (Kruger and Burrus). For example, while Duke Undergraduates appear to be consistently optimistic about their financial futures, it may be the case that the student population at a less prestigious state school might contain a much higher proportion of individuals who would be pessimistic. This raises the question of how my method copes in a domain characterized by a mix of individuals holding both optimistic and pessimistic views. Here I can not fall back on an idiosyncratic ideal. I demonstrated in study 10 that the “in a worst case” variant of the ideal world question could be used to attenuate pessimistic biases (e.g., below average effects) in domains where such biases are the norm. However, if some domains tend to be characterized by both significant populations of individuals holding both optimistic and pessimistic views then any debiasing effort essentially needs to know which
individual is which in order to ask the right question. As such, an intriguing possible route to universally encourage individuals to test more realistic hypotheses in such scenarios may be to have them consider both best and worst case scenarios before forming an estimate. Of note, such a method may actually have the potential to generalize across all domains. Future research could usefully explore this.

Future Research

In addition to further exploring how variations of the ideal world question affect behavioral predictions, there are a number of different directions in which this research might proceed. As far as the selective accessibility model, several avenues seem worthy of exploration: First, studies 9 and 10 raised the question as to whether selective hypothesis testing might contribute to biases beyond unrealistic optimism per se. While these studies focused on biased perceptions of trait standing, it is conceivable that that this technique might help with other biases that may in fact be characterized by selective testing over an overly positive hypothesis. The hindsight bias (Arkes et al. 1981) and the illusion of control (Langer 1975) may have particular promise in this regard.

As far as marketing is concerned I think several possible extensions are particularly interesting: First, I currently have only limited evidence that the decreased reliance of idealistic thinking engendered by the two-question method genuinely
motivates compensatory strivings to behave differently, and there is a lot more research that could be done in this area. For example, one could examine whether motivation toward important (and socially desirable behaviors) such as recycling and charitable giving can be increased by having individuals more realistically address their own and/or society’s shortcomings in these areas.

Second, managers interested in estimating trial rates for new products routinely encounter consumers who overstate how likely they are to purchase new products. To get better estimates of trial rates, market researchers typically use test markets or simulated test markets, both of which are very costly. If the two-question approach can produce better self-reported trial data, then it could also prove to be a useful tool to assist in demand estimation.

Conclusion

Consistent with theories of self determination I find that the key to helping consumers make realistic predictions is getting them to do so of their own accord. Namely, my studies reveal that realistic consumer predictions do not follow from requests to be realistic, but rather from an approach that works within the underlying process of selective hypothesis testing. In other words, the key to more realistic predictions of future behavior lies not in exhorting consumers to ignore the ideal, but in
getting them to acknowledge it. Those who did so (of their own accord) abandoned their idealistic expectations and embraced more realistic ones.
REFERENCES


BIOGRAPHY

BORN 28 August 1969, Epsom, Surrey, UK.

EDUCATION Ph.D. Duke University, Fuqua School of Business, Marketing, May 2008 (expected)

M.B.A. University of Chicago, 1999

B.Sc. Industrial Economics, Warwick University, Coventry, UK, 1990

PUBLICATIONS


PROFESSIONAL AFFILIATIONS

Association for Consumer Research

Society for Consumer Psychology