

Behaving Optimistically: How the (Un)Desirability of an Outcome Can Bias People's Preparations for It

JILLIAN O' ROURKE STUART^{1*}, PAUL D. WINDSCHITL¹, ANDREW R. SMITH² and AARON M. SCHERER³

¹University of Iowa, USA

²Appalachian State University, USA

³Center for Bioethics and Social Sciences in Medicine, University of Michigan, USA

ABSTRACT

Past research on the desirability bias and on bracing for bad news has focused on the potential influence of outcome desirability on people's stated expectations. The present studies examined its influence on behavior—that is, what is done in anticipation of, or preparation for, an uncertain outcome. In five studies, the desirability of possible outcomes for an event, which was uncertain and uncontrollable by the participant, was manipulated, and preparation behavior was measured. Study 1 used a hypothetical-events paradigm. Studies 2 and 3 involved a computer activity in which behavior was tracked on a trial-by-trial basis. In Studies 4 and 5, the uncertain event was the ending of a videotaped basketball game. Rather than exhibiting bracing or a reluctance to tempt fate, participants tended to behave in a manner consistent with an optimistic desirability bias. In a subset of studies, predictions and likelihood judgments were also solicited; the differential effects of outcome desirability on these measures are discussed. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS desirability bias; wishful thinking; optimism bias; preparedness; readiness; decision making; risk

People often face situations with uncertain outcomes that they cannot control but must do their best to anticipate. In these situations, decisions about how to act in preparation for key outcomes should depend on the likelihoods of the key outcomes—not on the desirability of those outcomes. Consider an example in which an employee has been informed that there is a chance of being sent to a distant city for a summer traineeship. Assume this employee has no say in the matter. The employee's decisions about how much to prepare for that possibility (e.g., by researching apartment availability or even placing a deposit on an apartment) should depend on the perceived likelihood of being sent on the traineeship—not on how attractive the employee finds the traineeship or its location.

In this paper, we address the question of whether people sometimes let the desirability of a key outcome influence their actions in anticipation of it—even when they cannot control that key outcome and desirability is therefore irrelevant. If desires bias preparations for an outcome, there are two possible directions of influence. There might be an optimistic desirability bias, whereby desire (dislike) for a key outcome increases (decreases) people's tendency to act like it will happen, or there might be a pessimistic desirability bias, whereby desire (dislike) for an outcome decreases (increases) our tendency to act like it will happen. While there are many studies addressing related issues (discussed later), we do not know of any published studies that experimentally manipulated outcome desirability to test for a biasing effect on preparation behavior. Here, we present five such studies.

One way in which either an optimistic or pessimistic bias in behavior might materialize is if the desirability of the outcome biases people's perceptions of the likelihood of

the outcome, by biasing the way people seek, interpret, or use evidence relevant to the outcome (for an extended discussion, see Krizan & Windschitl, 2007). If outcome desirability does tend to bias people's perceptions of likelihood in one direction or another, then it stands to reason that their behavior would be biased in the same direction. However, in studies that specifically examined the causal role of outcome desirability on subjective likelihood estimates, evidence as to whether outcome desirability inflates estimates (i.e., a *desirability bias/wishful thinking*), deflates estimates, or has little net impact has been mixed.¹ A key conclusion from a 2007 meta-analysis of studies that experimentally manipulated outcome desirability was that there is little overall evidence that outcome desirability impacts likelihood estimates in either a positive or negative direction (Krizan & Windschitl, 2007). Researchers who systematically attempted to find evidence of wishful thinking on likelihood judgments had—after little success—aptly described the wishful thinking effect as “elusive” (Bar-Hillel & Budescu, 1995). Although some studies since this 2007 meta-analysis have revealed significant desirability biases/wishful thinking (e.g., Lench, 2009; Windschitl, Scherer, Smith, & Rose, 2013), others have revealed the opposite (e.g., Bilgin, 2012; Vosgerau, 2010). This has led some researchers to question whether the desirability bias exists (Harris & Hahn, 2011; Vosgerau, 2010; see also Bar-Hillel & Budescu, 1995).

Despite the fact that the evidence that manipulations of outcome desirability influence likelihood estimates is mixed, our main prediction in the current studies was that such manipulations would tend to affect preparation behavior in

*Correspondence to: Jillian O'Rourke Stuart, Department of Psychology, University of Iowa, Iowa City, IA, USA. E-mail: jillian-stuart@uiowa.edu

¹Most studies on optimistic biases do not specifically attempt to isolate the causal role of outcome desirability. They leave outcome desirability as one of multiple reasons why a stated expectation for a desired (undesired) outcome is higher (lower) than an objective standard (for discussion, see Windschitl & Stuart, in press; see also Krizan & Windschitl, 2007, 2009).

an optimistic direction. Inspiration for this prediction comes, in part, from a set of studies showing that, when expectations are solicited in the form of dichotomous outcome predictions rather than in the form of likelihood estimates, manipulations of outcome desirability do have a significant impact (for a review, see Krizan & Windschitl, 2007). Many of these studies utilized what is called the marked-card paradigm. In that paradigm, initiated by Marks (1951), participants make predictions about whether a card—drawn from a deck with known numbers of marked and unmarked cards—will be marked. A monetary manipulation makes drawing a marked card desirable, neutral, or undesirable. When the proportion of marked cards is at or near 50%, participants are more likely to predict a marked card when drawing marked cards are desirable rather than not (e.g., Irwin, 1953; Irwin & Snodgrass, 1966; Lench & Ditto, 2005; Pruitt & Hoge, 1965; for meta-analysis, see Krizan & Windschitl, 2007). This effect of desirability on outcome predictions has been detected even when the same manipulations failed to produce detectable effects on likelihood judgments within the same studies (Price & Marquez, 2005; Windschitl, Smith, Rose, & Krizan, 2010).

Our prediction about behavior extends from an account suggested by Windschitl et al. (2010) for why manipulations of outcome desirability influence predictions—even when they do not influence likelihood estimates. This *biased-guessing account* posits that in situations where participants are forced to make a subjectively arbitrary prediction (e.g., a dichotomous guess), they will typically guess in an optimistic rather than pessimistic direction. This is particularly relevant in situations where there is no definitive information as to how one should respond, such as in the marked card paradigm when there was an equal or roughly equal number of marked and unmarked cards. Likelihood judgments may be less vulnerable to the influence of desirability because participants initiate a scaling process that, under most circumstances, does not involve guessing. The judgments are scaled based on perceptions of evidential support (for an example model, see Tversky & Koehler, 1994), and desirability manipulations appear to have less impact on evidence perceptions than on guessing processes (see Windschitl et al., 2010, for more information). Here, we propose that desires will influence behavior in a manner similar to how, in previous work, they influenced predictions. Both a prediction and a behavior involve a choice rather than merely a judgment; they both involve orienting toward a particular outcome as the one that will presumably happen. Under some degree of evidence ambiguity or balance, this orienting process can involve guesswork and is susceptible to the influence of outcome desirability. Of course, if outcome desirability also influences perceptions of evidence, this could contribute to an effect on behavior.

A set of studies that used the marked-card paradigm but solicited bets rather than predictions lends initial yet limited empirical support for our prediction that behavior would be optimistically biased by outcome desirability (e.g., Irwin & Snodgrass, 1966; Morlock & Hertz, 1964). Participants bet more on a marked card when drawing a marked card was desirable rather than not. Obviously, placing a bet is a form

of behavior, so these studies are relevant to our prediction. However, measuring bets is only one way of operationalizing behavior, and it may have a peculiar property in the context of the marked-card paradigm, where both bets and desire are defined monetarily. Betting on the undesirable outcome would clearly be hedging or betting against the self. It would ensure the gain of a desirable draw would be cut by the loss from the misplaced bet. The present research does not involve bets nor monetary manipulations. We measure behavior that is better characterized as preparing for, or anticipating, an outcome, rather than wagering about an outcome.

In opposition to our main prediction, it is also possible that outcome desirability would have no effect on behavior or even a pessimistic effect. In our studies, the outcomes were impending. As an outcome draws near, people may become less comfortable taking an optimistic stance toward an uncertain event, and they might even embrace a pessimistic one (Gilovich, Kerr, & Medvec, 1993; Shepperd, Ouellette, & Fernandez, 1996; Shepperd, Findley-Klein, Kwavnick, Walker, & Perez, 2000; Tyler & Rosier, 2009; for a review, see Sweeny & Krizan, 2013). Specifically, people may strategically adopt a position—that is, a subjective probability, a prediction, or a behavior—that is pessimistic because it would mitigate the added unpleasantness of negative outcomes that are unexpected rather than expected (Shepperd et al., 1996; Sweeny, Carroll, & Shepperd, 2006). The concept of an asymmetric loss function is also relevant here. Work on asymmetric loss functions suggests that people implicitly or explicitly know that the harm of being overly optimistic is typically greater than the harm of being overly pessimistic—leading to a more pessimistic stance (Haselton & Nettle, 2006; Weber, 1994; Weber & Hilton, 1990).

An entirely different reason why people might behave in a way that appears pessimistic concerns the notion of tempting fate. *Tempting fate* is the superstitious belief that it is bad luck to behave with hubris or do something that is interpreted as involving needless risk (e.g., a college applicant wearing a sweatshirt from her top-choice university prior to being offered admission; Risen & Gilovich, 2008). Some people have the intuition that tempting fate will actually decrease the likelihood of a desired outcome or increase the likelihood of an undesired outcome (Miller & Taylor, 1995; Risen & Gilovich, 2007, 2008; Swirsky, Fernback, & Sloman, 2011; Van Wolferen, Inbar, & Zeelenberg, 2013; see also Kogut & Ritov, 2011). Referring to our opening example about the employee who might be assigned to a summer traineeship, if that employee was impacted by tempting-fate beliefs, he or she might be overly cautious and avoid looking at apartments precisely because he or she desires the position and does not want to lessen the likelihood of getting it. Although various studies have tested when and why people sometimes believe that tempting fate reduces the likelihood of a desired outcome, no published studies involve the type of tests we conducted for this paper. The typical study on tempting fate presents people with a scenario in which a protagonist did or did not “tempt fate,” and the dependent measures ask people to estimate the likelihood of a desired outcome given the protagonist’s actions (e.g., Risen &

Gilovich, 2008; Swirsky et al., 2011). In other words, the studies did not measure people's behavior and therefore did not gauge whether behavior tended to be optimistically or pessimistically biased. Our studies were specifically designed to gauge this.

Overview of present studies and reporting standards

We used three paradigms to conduct a series of studies. Study 1 assessed hypothetical behavior in an everyday scenario. Studies 2 and 3 measured actual behaviors in the context of a novel computer game. Studies 4 and 5 also assessed actual behaviors, but this time in response to the uncertain outcome of a videotaped basketball game. In addition to behavior measures, Studies 3–5 also included perceived likelihood measures. For each study, we set a sample-size target in advance and stopped data collection based on this target and various logistical factors (e.g., participant sign-ups and experimenter availability). All measures, conditions, and data exclusions are reported. No covariates were used.

STUDY 1

Participants in Study 1 read a hypothetical scenario about a potential upcoming event, described as either desirable or undesirable. They were then asked if they would engage in a behavior that involved preparing for the event or if they would do nothing. A tempting fate hypothesis would predict that for a desired event, people would hesitate to prepare for it, so as to not trigger its undoing. A pessimism hypothesis might suggest that for a desired event, people would tend to view the event as unlikely and therefore not prepare for it. Contrary to these predictions, we expected that people would tend to act optimistically. That is, they would be more likely to prepare for a desired event than an undesired one.

Method

Participants

The participants were 273 Amazon Mechanical Turk workers paid a small fee for their participation (62.6% male; 13 did not provide gender; M age = 29.67, 10 did not provide age).

Procedure and materials

Participants read a hypothetical scenario about the potential installation of a public pool in the area behind their house (see Appendix A for full scenario). They were told that the determining factor for whether the pool would be located in that area was the results of foundation testing that had not yet been conducted and whose results were described as uncertain. Half of the participants were told to imagine that they really wanted this pool. They also read reasons why it would be personally desirable (e.g., easy access for your family), and they viewed a picture of an attractive public pool. The other half of participants were told to imagine that they really did not want this pool. They read reasons why it would be

personally undesirable (e.g., much noise and traffic associated with public pools), and they viewed a picture of a less attractive public pool.² In addition, all participants learned that if the pool is built, the city would require them to make some landscaping changes to the back end of their yard. If they opt to make the changes now, before the results of the foundation testing are confirmed, they could obtain a 70% discount on costs by using a crew that was currently employed by a neighbor. Participants were asked "Given the discount, would you go ahead and have the changes to your property made even though you can't be sure about whether the pool would go in?" (1 = *definitely would not*, 2 = *probably would not*, 3 = *probably would*, 4 = *definitely would*). They then completed one open-ended question regarding why they answered the way they did to the previous question, a set of manipulation checks, and demographics.

Results

Three manipulation check items assessed how much the scenario made it seem like having a pool would be desirable and how appealing living near a public pool would be generally. These three items were significantly correlated ($r_s > .36$, $p_s < .001$), and we combined them to form a composite measure. Analysis of the composite measure confirmed that we successfully manipulated participants' desire for the pool (desirable $M = 5.17$, $SD = 1.24$; undesirable $M = 2.24$, $SD = 1.23$), $t(260) = 19.22$, $p < .001$, $d = 2.37$. Analyzing each measure independently yielded the same results.

Regarding participants' behavior, both Mann–Whitney and t -tests revealed the presence of a desirability bias, $U = 7102.00$, $p < .001$; $t(265.21) = 3.77$, $p < .001$, $d = 0.46$. As predicted, participants who read that they desired the pool were more likely to indicate that they would make the landscaping changes ($Mdn = 3$, $M = 2.94$, $SD = 0.72$) than did participants who read that they did not desire the pool ($Mdn = 2$, $M = 2.58$, $SD = 0.84$).

Discussion

Study 1 demonstrated that, when people have an option between preparing for a potential outcome or not, the manipulated desirability of that outcome biases their preparation choice. In this study, the direction of the bias fit optimism, rather than pessimism or a tempting fate concern.

Concurrently with Study 1, we conducted a related study with a separate sample of 264 participants from Mechanical Turk. We used a similar scenario, except the pool was a dog park (which, depending on results of environmental-impact testing, might be built), and we did not manipulate desirability (see Appendix B for full scenario). Instead, we anticipated that the perceived desirability of having a dog park constructed near one's home would vary substantially across people in our sample, and we therefore simply

²Because any particular pool picture may have unforeseen idiosyncrasies, we used three pictures within each condition, but a given participant only saw one. There was no effect of this counterbalancing factor ($p_s > .39$), so it will not be discussed further.

measured subjective desirability on seven point scales. Dog-park desirability did indeed vary substantially. More important, the primary analysis produced a result consistent with Study 1 and a desirability bias. The more that participants desired a dog park, the more inclined they were to initiate the landscaping changes to their yard that would only be required if the dog park was built ($r = .49$, $p < .001$). Again, the building of the dog park depended on the results of environmental-impact testing. For this uncontrollable event, participants were more interested in preparing for the event if they saw it as desirable rather than undesirable.

STUDY 2

For Study 2, we developed a computer-game paradigm that allowed us to test actual behavior, rather than hypothetical behavior. During game trials, participants needed to prepare for an uncertain event that had two possible outcomes, one of which was more desirable than the other.

The details of the game are described in the "Method" section, so we provide only a brief overview here. Each trial involved a screen split down the middle, with many circles on each side (Figure 1). Participants knew that an unpredictable "slider" was about to appear and would travel down either the left or right side of the screen. Based on the desirability manipulation, participants hoped the slider would travel down a particular side, because this would yield a significant point gain. However, participants also knew that they had a limited time to "prepare" circles for the slider (by clicking on them). They knew that, irrespective of whether the slider went left or right, they would earn points for every "prepared" circle that was hit by the slider. Circles differed in size, which affected how easy it was for participants to quickly prepare them with an accurately placed mouse click. Given this situation, the optimal strategy for the short circle-preparation period was to start with the large/easy circles, irrespective of what side the participant hoped the slider would travel down. If, however, participants exhibited an optimistic desirability bias, they would tend to prepare circles

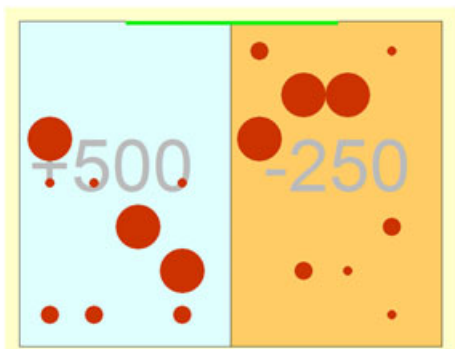


Figure 1. An example screenshot of the computer game from Study 2. The slider is the thin green bar depicted at the top of the screen. After a brief period in which participants could click on circles to "prepare" them, the slider moved to either the left or right and slid down that side of the screen. Participants earned or lost points depending on the side of the screen the slider went, but they also gained points for every prepared circle that the slider hit

on the desirable side, perhaps at the cost of preparation efficiency. The other alternative, of course, was that participants would prepare more on the undesired side, reflecting a pessimistic perspective or a form of hedging (i.e., ensuring that if they miss out of a reward because the slider failed to travel to the desired side, they would at least earn points when the slider hits the circles they prepared on the undesired side).

Method

Participants and design

Fifty-five undergraduate students at the University of Iowa participated. The design was a 2 (counterbalance: desirable side on the right or left) \times 5 (block: 1, 2, 3, 4, 5) mixed factorial. Both desirable and undesirable sides were present on each trial.³ Each block had 10 trials. The counterbalance factor did not produce main effects or interactions and will not be discussed further. Two exploratory blocks of trials also appeared after the first five (see details in the "Exploratory blocks and additional measures" section).

Procedures

The goal of the participants was to score points so they would be eligible for a candy prize. They went through four practice trials with extensive feedback to ensure they understood the two ways in which they could win and lose points. Next, they completed 50 real trials broken up into five blocks of 10 trials. Each trial had the same parts. Figure 1 depicts what participants saw immediately after clicking "go" to start a trial. There were nine red circles randomly placed on each side of the screen (three small, medium, and large). There was "+500" printed on the background of one side and "-250" on the other (counterbalanced). Participants had 4.5 seconds to prepare as many circles as possible by clicking on them, which turned them yellow. When the time expired, a rectangular bar called a "slider" appeared at the top center of the screen. It then immediately slid to the left or right side of the screen (randomly determined within blocks) and then traveled down that side. As the slider descended over the circles on that side of the screen, it turned all the "prepared" circles from yellow to green. The amount of points lost/gained on the trials was immediately calculated, with explicit labeling of how the points were lost/gained. Then the next trial began.

About points

From the initial instructions and extensive feedback in practice and test trials, participants were well informed about the two sources of points. First, participants either gained 500 points or lost 250 points as a function of whether the slider picked the +500 or -250 side. This served as our desirability manipulation. Second, participants received 25 points for

³This may be considered as another within-subject factor, but because of the way we scored the dependent measure (creating an index that compares behavior on the two sides), it makes more sense to leave this factor out of the factorial design statement.

each circle that they had prepared on the side that was selected by the slider.

Exploratory blocks and additional measures

When designing this study, we expected that most participants would be significantly fatigued by the end of 50 trials. Nevertheless, for exploratory purposes, we added two additional blocks of trials. There were two types of blocks, presented in counterbalanced order. In one block, participants gave 2 likelihood estimates at the start of every trial—the likelihoods that the slider would chose the left and right sides (adding to 100%). At the start of every trial in the other block, participants predicted whether the slider would choose the left or right side.

Participants then completed the Life Orientation Test-Revised (LOT-R) scale (Scheier, Carver, & Bridges, 1994), the Need for Cognition Scale (Cacioppo, Petty, & Kao, 1984), and the Numeracy Scale (Lipkus, Samsa, & Rimer, 2001). Scores from these scales were not significantly correlated with our key measures and will not be discussed further. Before being debriefed and receiving candy, participants completed an exit questionnaire with items about the tendencies of the slider, self-perceived prediction ability, performance and strategy changes across the course of the session, interest and effort in the task, and speculation about the purpose of the task. None of these items yielded results that were important to discuss here.

Results and discussion

Within the main five blocks of trials, participants were able to prepare an average of 8.79 ($SD=0.80$) circles per trial. In order to assess whether participants were optimistically biased in their preparation of circles, we calculated what we will call the *desirability-bias index*, namely, the percentage of prepared circles that were from the desired rather than the undesired side (for a given trial). This index is plotted by block in Figure 2.

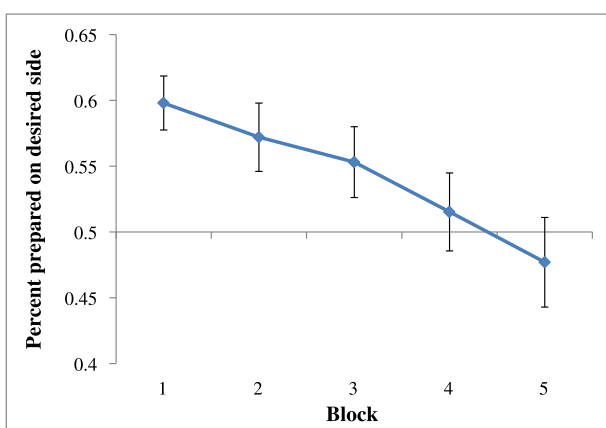


Figure 2. Desirability-bias index by block (Study 2). Error bars represent ± 1 standard error

In the main blocks, the overall average index was 54.3% ($SD=16.7\%$), which represents a marginally significant tendency to prepare more circles on the desired side, $t(54)=1.91$, $p=.06$, $d=0.26$. On closer analysis, the bias was robust in early blocks and tapered off in later blocks. More specifically, an average of 59.8% ($SD=15.2\%$) of the circles prepared by participants in Block 1 were from the desirable side, which was significantly above 50%, $t(54)=4.78$, $p<.001$, $d=0.64$. The desirability-bias index was also significantly above 50% in Blocks 2 ($t(54)=2.77$, $p=.008$, $d=0.37$) and 3 ($t(54)=1.97$, $p=.05$, $d=0.27$), but not in Block 4 ($t(54)=0.52$, $p=.61$, $d=0.07$) or 5 ($t(54)=-0.67$, $p=.50$, $d=0.09$). An analysis of variance (ANOVA) examining the desirability-bias index across the five blocks revealed a significant main effect of block, $F(2.36, 127.70)=6.91$, $p<.001$, $\eta_p^2=0.11$, indicating that there was a significant decrease in the optimistic desirability bias over time. Other research on the desirability bias has found similar learning effects where the bias decreases over multiple trials (Bar-Hillel & Budescu, 1995, Study 4), presumably because participants begin to understand that their biased strategy is not the most effective at maximizing gains.

There were also results consistent with the idea that exhibiting a desirability bias in behavior was suboptimal. In preliminary steps for the analyses, we computed an index of the average size of the prepared circles (where small, medium, and larger circles were coded as 1, 2, and 3, respectively). The average was above 2 ($M=2.37$, $SD=0.17$, $p<.001$), reflecting an overall tendency to click on larger (52%) rather than medium (33%) or smaller (15%) circles. We also tallied the number of circles successfully prepared. Critically, among the participants whose desirability-bias index was in the optimistic direction (i.e., directionally above 50%; $n=34$), there was a significant inverse correlation between the desirability-bias index and performance (i.e., number of circles prepared; $r=-.41$, $p=.02$). It appears that being more optimistically biased came with a cost. Other correlations shed light on why. An inverse correlation between the desirability-bias index and average-size index ($r=-.53$, $p=.001$) suggests that being biased led to less of a strategic focus on large circles. And, a positive correlation between the average-size index and number of circles prepared ($r=.46$, $p=.006$) is consistent with the notion that participants who failed to focus on clicking mostly large circles (i.e., the easy-to-click ones) did not perform well overall.

Results from the exploratory blocks

Given that the desirability bias on behavior had diminished and was not significant by Blocks 4 and 5, it is not surprising that the desirability bias on behavior was also not significant in the exploratory trials of Blocks 6 and 7 ($p=.68$ for the block that included likelihood judgments; $p=.41$ for the block that included outcome predictions). Average likelihood judgments about whether the slider would go to a given side were not significantly affected by the desirability of the side ($p=.84$). The effect of desirability on outcome predictions was significant ($p=.05$); the mean rate at which participants predicted the desirable side was 56% ($SD=22.6$).

STUDY 3

In Study 3, we used largely the same game as in Study 2, but with four key changes. First, we solicited likelihood judgments in addition to behaviors (in counterbalanced order). In the introduction of this paper, we discussed that the effect of outcome desirability on behavior might be robust even when participants' subjective likelihood estimates are not affected or are much less affected. This was inspired by the *biased-guessing account* for why manipulations of outcome desirability influenced predictions even when they did not influence likelihood estimates (Windschitl et al., 2010). The account previously posited that, in situations where people must make a subjectively arbitrary prediction (e.g., a dichotomous guess), they will typically guess in an optimistic rather than pessimistic direction. When making likelihood judgments, guessing is usually not a factor; the judgments may be less vulnerable to the influence of desirability because participants initiate a scaling process based on perceptions of evidence (Bilgin, 2012; Windschitl et al., 2010). Given this account and empirical work already discussed, we expected there to be little if any effect of desirability on likelihood judgments, but we did expect a replication of the effect of desirability on behavior.⁴

The three remaining changes extended the generalizability of the findings from Study 2. For the second change, we made the behavioral measure a dichotomous choice. Participants now had to choose between preparing circles only on the desirable side or only on the undesirable side, rather than being able to prepare circles on both sides of the screen. For the third change, instead of using point gains and losses to manipulate the desirability of the slider going to the left or right sides, we told participants that depending on which side the slider chose, they would be required to either solve a difficult math problem or read a short joke (which informal pilot testing suggested were undesirable and desirable, respectively). For the fourth change, participants in Study 2 were given more information relevant to the likelihood that the slider would go left or right on a given trial. They saw a two-color grid before each trial (Figure 3). Each color represented a side, and the direction of the slider on that trial was said to be based on the color of a randomly selected square from the grid on that trial. To reinforce the meanings of the grid squares, we also inserted a small "M" in each square containing the color associated with math problems and a "J" in each square containing the color associated with jokes (Figure 3). On critical trials, the grid colors always had proportions of 50/50, yet we assumed there would be trial-by-trial variation in participants' perceptions of the actual proportions of these grids. We wondered whether the presence of the grid would cause participants to ignore

desirability and focus on determining which side was more likely in a given trial. This might happen because the presence of objective evidence could put participants in a more analytical mindset. If so, this would place a rather restrictive boundary condition on the optimistic behavior we had observed in Study 2.

Method*Participants and design*

The participants were 137 undergraduates from the University of Iowa. The study design consisted of a 2 (order: choice first, likelihood judgment first) × 2 (first location counterbalance: desirable side on the left or right) × 2 (block: 1, 2) mixed factorial with only the last factor as within subjects.

Description of game

The specific differences between Studies 2 and 3 are in conjunction with the four differences we have already discussed earlier.

Two practice trials were included to orient participants to the newly added grids. Instructions emphasized the utility of the grid: "You will see a grid that will help you predict which side the slider..." Four additional practice trials oriented participants to why the background of one side of the screen said "joke" and the other "math," and why the grid squares contained "J" or "M." Participants viewed two math problems that we thought most students would find aversive and two jokes that were generally amusing.

After the instruction and practice phase, participants engaged in 16 trials (split into two blocks of eight) of which 12 were critical trials and four were filler trials. Each trial began with the display of a two-color grid. Again, the grid essentially revealed the likelihood of each possible outcome—the slider going to the math side or the joke side. The color pairs used for the grid and sides changed from trial to trial, to accentuate the independence of the trials. In addition to counterbalancing whether the left or right side was initially associated with jokes/math, this initial pairing was switched after the first block of eight trials. In the grids for the critical trials, the proportion of squares representing the math and jokes sides was always 50/50. For the filler trials, the proportions were either 30/70 or 70/30. We included these filler trials so that participants would not assume that all grids have 50/50 proportions.⁵

After seeing the grid for a given trial, participants made either a choice and then a likelihood judgment, or a likelihood judgment and then a choice. As outlined earlier, we changed behavior to a categorical choice. This meant that now participants picked one side to prepare and then were given 4.5 seconds to click on circles only on the side they had chosen. For the likelihood judgments, participants were asked how likely it was that the computer would pick one side or the

⁴Given our use of counterbalancing, we were also poised to examine order effects. It seemed possible that the magnitude of the desirability bias on whatever measure was asked first could carry over and influence responses to the second measure. For example, perhaps the act of providing a scaled likelihood judgment would de-bias participants and subsequently rein in their behavior (see Bouts & Van Avermaet, 1992, for an analogous example in which the biasing influence of familiarity was observed on betting decisions, except when likelihood judgments were solicited first).

⁵Not surprisingly, there was no evidence of a desirability bias in the filler trials; the 30/70 proportions in the filler trials gave participants a relatively strong indication as to the likely direction of the slider (also Windschitl et al., 2010).

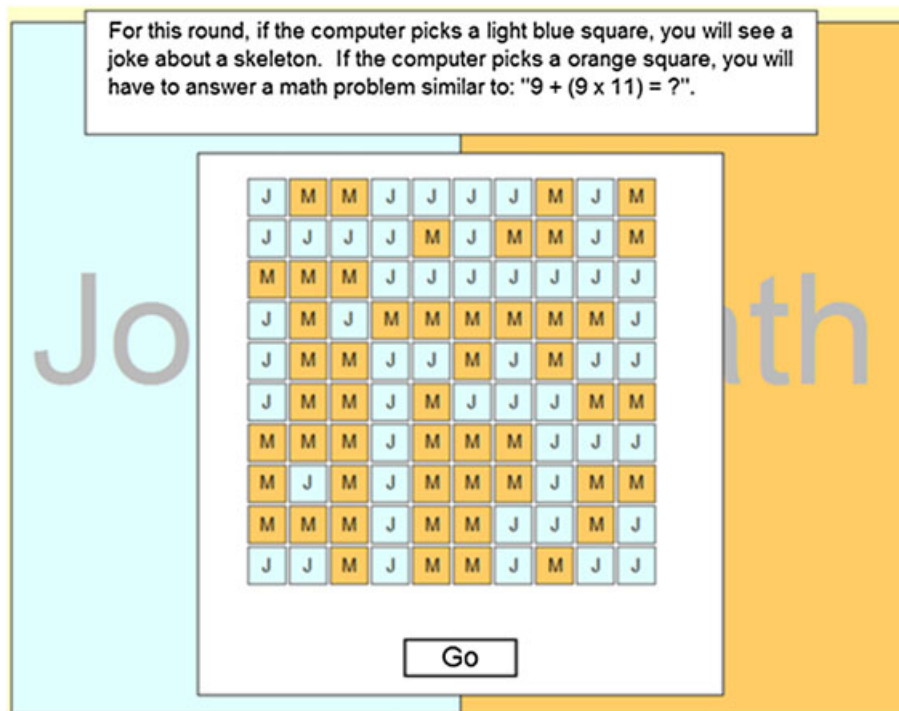


Figure 3. Example of the matrix and instructions participants saw before a trial in Study 3

other. The 5-point response scale ranged from *very likely to go [color of right side]* to *very likely to go [color of left side]* and was presented vertically on the screen. After providing their responses on a given trial, participants proceeded to the next trial without seeing which side the slider selected or the number of points earned/lost in a given trial. This feedback was provided later, and all participants saw eight jokes and answered eight math problems at the end of the study.

Post-game measures

To more closely examine the impact of specific color pairings on the perceived proportions on the grids, we added six testing trials after the game. In each trial, participants viewed one of the grids and color pairings. Each square of a given color also contained an “M” or “J” (fully counterbalanced). After 4 seconds, the grid disappeared, and participants were asked to indicate which color was more frequent on the grid.

Participants also completed the LOT-R scale (Scheier et al., 1994), the Numeracy scale (Lipkus et al., 2001), the Behavioral Inhibition and Activation Scale (BIS/BAS; Carver & White, 1994), and the Belief in Good Luck Scale (Darke & Freedman, 1997). Scores from these scales were not significantly correlated with our key measures and will not be discussed further.

Results

Behavior

Overall, there was a significant desirability bias for behavior—the rate at which participants chose the desirable side to prepare on critical trials was 55.1% ($SD=12.8\%$), $t(135)=4.62$, $p < .001$, $d=0.40$. A 2 (order) \times 2 (counterbalance) \times 2

(block) ANOVA on the desirability-bias composites (calculated per block) revealed that the desirability bias did not significantly differ as a function of order ($p=.67$). The composites for the bias were significantly above 50% regardless of whether choices were solicited before ($M=54.7\%$, $SD=11.2\%$) or after ($M=55.5\%$, $SD=14.4\%$) likelihood judgments (both $ps < .01$). The 2 \times 2 \times 2 ANOVA also revealed no significant effect of block ($p=.18$), nor were there any interactions ($ps > .37$). The only significant main effect was an unimportant effect of the location counterbalancing condition (i.e., whether jokes first appeared on the left or right; $p=.02$).

Likelihood judgments

We rescored the likelihood judgments such that 5 reflected *high certainty* and 1 *low certainty* about the desirable outcome. Despite the fact that the critical trials all involved grids containing 50/50 proportions, the middle response option on the likelihood scale was selected in only 49% of trials overall. In other words, there was notable variability in how people viewed the likelihoods of outcomes when they examined the grids.

More important was whether there was a desirability bias. The effect was significant, which was not expected, but it was small. The overall mean on the critical trials ($M=3.05$, $SD=0.21$) was slightly greater than 3—the midpoint of the scale, $t(135)=2.94$, $p=.004$, $d=0.24$. As with behavior, a 2 (order) \times 2 (counterbalance) \times 2 (block) ANOVA revealed that likelihood judgments did not significantly differ as a function of order ($p=.99$) or block ($p=.11$). The only significant effect was again an unimportant main effect for location counterbalancing ($p=.01$).

Behavior versus likelihood judgments

Not surprisingly, participants' choices and likelihood judgments were associated. For each participant, we calculated the Fisher-transformed idiographic correlation across trials between likelihood judgments and choices (both scored so higher values favor the desirable side). The correlations tended to be high and significantly greater than zero regardless of whether choices ($M_r=0.66$; $t(64)=18.02$, $p < .001$) or likelihood judgments ($M_r=0.66$, $t(60)=17.67$, $p < .001$) were solicited first.

Various analyses provided conflicting conclusions about whether the effect of desire on behavior should be considered at least partially independent of its effect on likelihood judgment. To directly compare the magnitude of the desirability bias on behavior versus on likelihood judgments, we first calculated, for each participant, the proportion of times the left side was picked when it was desirable and when it was undesirable. We standardized those scores. Next, we calculated for each participant the mean likelihood estimate for when the left side was desirable and for when it was undesirable (with scores coded such that high values meant high perceived likelihood of the left side). We standardized these scores. These four sets of scores were submitted to a 2 (desirability) \times 2 (type of measure) repeated-measures ANOVA. A significant interaction revealed that the effect of desirability was larger on behavior than on likelihood judgments, $F(1, 135)=4.75$, $p = .03$, $\eta_p^2 = 0.03$.

We also conducted mediation analyses using participant-level data and the guidelines defined by Judd, Kenny, and McClelland (2001) for within-subject designs. Participants' tendency to give higher likelihood estimates for a given side when the side was desired (vs. not) was strongly correlated with their tendency to prepare for that side when it was desired (vs. not), $r = .60$, $p < .001$. Given this tight relationship, it is not surprising that the mediation analyses returned results that were consistent with the idea that the effect of desirability on likelihood judgment partially mediated the effect of desirability on behavior. However, the causal interpretation of the pattern is not clear (e.g., a third variable explanation is that participants who most cared about outcome desirability were the most likely to exhibit a strong desirability effect on both their likelihood judgments and behaviors).

Finally, we also found that when examining only trials in which a given participant had responded with the middle likelihood judgment, participants chose to prepare the desired side at an average rate of 56%, which is significantly greater than 50%, $t(129)=3.13$, $p = .002$, $d=0.27$. That is, even on trials in which likelihood judgments were all the same (and neutral), a desirability bias in behavior was evident.

Influence of color pairings and letters

The analyses reported earlier collapsed across a counterbalancing of which color in a given pair was associated with the desired side/outcome. This ensured that color pairings were orthogonal to the results of the desire manipulations. Nonetheless, we briefly report here on an interesting effect of color revealed by examining results at the individual trial level. In short, it appears that lighter colors in the grid

were more visually salient. On seven of 12 critical trials, one color from a pair (i.e., the side associated with that color) was predicted significantly more than the other color and was given a higher likelihood estimate ($ps < .05$). Also, on five of the six post-game trials, there was a significant bias toward one color over the other ($ps < .05$). In virtually all these cases of imbalance, the lighter color in the grid was predicted or seemed more likely/frequent than the darker colors. Again, these findings are orthogonal to our main findings. Another notable finding from the post-game trials was that colors were judged as more frequent when their squares contained the letter "J" (chosen on 56% of the post-game trials) rather than "M." We highly doubt that this, itself, is a visual salience effect. Instead, it likely reflects carryover from the previous 16 trials they experienced, in which "J" was always signifying the potential for a more desirable outcome. In a similar study that used grids but did not display Ms or Js (and involved no likelihood judgments), the basic desirability bias on behavior was replicated, providing additional reason to believe that the letters are not critical to the desirability bias observed in this study.

Discussion

The results of Study 3 extend the generalizability of the findings from Study 2 in at least three ways. First, even when participants were given concrete information about the likelihood of desirable outcomes (via the grids), they still exhibited a significant desirability bias in their behavior. Second, we used jokes and math problems to manipulate desirability rather than points gained or lost. This change addresses a potential criticism of Study 2, which is that the reward earned by clicking on the right circles is the same currency as the desirability manipulation (i.e., the currency is points for both). Third, the desirability bias replicated with a dichotomous measure of behavior. In this study, we also solicited likelihood judgments, counterbalanced with measures of behavior. As expected, the effect of desire on the behavior measure was stronger than that on the likelihood measure, but the difference was smaller than anticipated. The effect of desire on likelihood judgments was significant. This issue is again addressed in the next two studies.

STUDY 4

Studies 2 and 3 assessed real behavior, rather than hypothetical behavior, to test for a desirability bias. However, the target events and other aspects of the computer game used in those studies (e.g., the repeated measures) could be considered abstract or artificial. In Studies 4 and 5, we again examined the effect of desirability on real rather than hypothetical behavior, but with a paradigm in which there was only one event being considered, the evidence regarding outcomes of the event was richer and less contrived than grid colors, and the preparation behavior being measured was more concrete.

The event that was used was a videotaped college basketball game. Before participants watched footage from the

game, we introduced a desirability manipulation. Namely, participants were randomly assigned to be associated with one team or the other; if “their” team won, they would receive a candy bar. Prior to finding out how the game ended, participants also learned that they would soon be doing one of two versions of an unrelated task—depending on the game outcome. The unrelated task was timed and involved generating poetic sentences related to a pool theme or park theme. As our main dependent variable, we gave participants a chance to prepare a list of rhyme words associated with one of the two themes—a list that might be useful for later generating the poetic sentences. Optimally, a participant should prepare rhyme words for the theme that seems most likely (given the state of the basketball game) or perhaps for the theme that seems more difficult. However, the outcome of the game was uncertain, and the two themes were designed to seem roughly equally difficult. In short, we set up a situation that we thought might be conducive to detecting a biasing influence of outcome desirability. We predicted that participants would be biased toward preparing as if the desired team was going to win (i.e., they would prepare for the rhyming task that would be relevant if their team won).

Participants were also asked to estimate the likelihood of the two possible game outcomes—either before or after deciding which rhyming task to prepare for. Soliciting these likelihood judgments provided another opportunity to compare the effect of desirability on behavior versus likelihood judgment. Based on the mixed literature regarding the effect of desirability on likelihood judgment, we did not expect a robust effect on likelihood judgments in this study. Dichotomous predictions about the game outcome were also solicited. For reasons discussed earlier, we expected predictions to align better with behavior than with likelihood judgments.

Method

Participants and design

Ninety-seven undergraduate students at the University of Iowa participated in a 2 (desired team: red, white) × 2 (order: behavior first, likelihood judgments first) between-subjects design.

Procedure, materials, and dependent measures

Participants were told that they would be watching a videotaped basketball game and that one of the teams would be designated as “their” team. If “their” team won the game, they would receive a candy bar. In view of each participant, the experimenter flipped a coin to assign the participant to either the red-jersey team (if heads) or the white-jersey team (if tails). Participants then watched footage from the last several minutes of a game from the 2010 tournament of the Women’s Northeast College Basketball Conference. We selected this particular game because it was unlikely that any of our participants would know about the outcome and because the score was very close in the last few seconds. One team had stormed back from being further behind, but it was unclear whether that team would overtake the other

before time ran out. The video was paused with 10 seconds left in the game, and participants were asked to complete the dependent measures in a counterbalanced order (either behavior, prediction, likelihood judgment or likelihood judgment, behavior, prediction).

For the behavioral measure, participants were informed that after the game was over, they would have to complete a timed task of writing poetic sentences that include rhyming words related to a specific theme. If the red team won, the required theme would be “afternoon at the park,” but if the white team won, the required theme would be “afternoon at the pool.” They then read:

Don’t worry, this isn’t a high-pressure thing, but you’ll do better at the poetic-sentences task if you prepare some rhyme words in advance. Right now, we are going to give you 2 minutes to prepare rhyme words for either the ‘Park’ or the ‘Pool’ theme... You have to pick a theme to prepare for even though you don’t know for sure what theme is relevant.

Then participants were prompted to click on one of two buttons to indicate the theme for which they would prepare: “the ‘pool’ theme (relevant if the white team happens to win)” or “the ‘park’ theme (relevant if the red team happens to win).”

For the prediction measure, participants were asked “What is your prediction about the outcome of the game? That is, which team do you think will win?”

For the likelihood measures, participants were asked separately about both the red and white teams’ likelihood of being the winner. They responded to each by selecting one of seven likelihood phrases on a scale from *extremely unlikely* to *extremely likely*.

Next, participants generated rhymes and then responded to a variety of questions about their experience viewing the game, how they made their decision, how desirable the reward was, whether they consider themselves an optimist or pessimist, their perceptions of the experiment, and if they could remember the team names. They completed the LOT-R (Scheier et al., 1994), Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988), and BIS/BAS (Carver & White, 1994). They watched the final 10 seconds of the game, followed by doing an untimed rhyming task and being debriefed.

Results and discussion

Preliminary analyses did not reveal noteworthy results regarding the additional questions and trait scales encountered by participants at the end of the session, so with one exception, those results are not discussed further. Table 1 displays means for the behavior, prediction, and perceived likelihood measures.

Behavior

A full 86.6% of participants chose to prepare for the rhyming task associated with their desired outcome—that is, with their teaming winning. Not only is this proportion significantly greater than 50% (binomial test, $p < .001$), it is also decidedly

Table 1. Percentages, means, and standard deviations for dependent measures in Study 4

	When red team winning was desired	When white team winning was desired	Overall sample
% Preparing for a red-team win	86.5	13.3	52.6
% Preparing for a white-team win	13.5	86.7	47.4
% Predicting a red-team win	84.6	15.6	52.6
% Predicting a white-team win	15.4	84.4	47.4
Mean likelihood estimate of red-team win (<i>SD</i>)	5.08 (0.81)	4.49 (1.31)	4.80 (1.11)
Mean likelihood estimate of white-team win (<i>SD</i>)	4.06 (1.04)	4.24 (1.32)	4.14 (1.17)

Note: Percentages are in pairs and add to 100. For example, 86.5% of the people who were told the red team was "their team" (i.e., red winning was desired) prepared for the rhyming task associated with the red team winning, while the remainder of those people (13.5%) prepared for the rhyming task associated with the white team winning.

stronger than the conceptually comparable rate observed in Study 3 (55.1%). This tendency to prepare as if the desired outcome would occur was not qualified by the counterbalancing of whether it was the red or white team that was made desirable, $\chi^2(1)=0.00$, $p=.99$. It was also not qualified by whether behavior was measured before (rate=84.3%) or after (rate=89.1%) likelihood judgments were solicited, $\chi^2(1)=0.48$, $p=.49$.

Predictions

The pattern of results for predictions was the same as for behavior. A full 84.5% of participants predicted that their desired team would win the game, significantly greater than 50% (binomial test, $p<.001$). Again, this tendency was not significantly qualified by the counterbalancing of whether it was the red or white team that was made desirable, $\chi^2(1)=0.001$, $p=.98$, nor whether predictions were measured before or after likelihood judgments were solicited, $\chi^2(1)=0.25$, $p=.62$.

Likelihood judgments

Recall that participants gave separate likelihood estimates of the red team and white team winning. Preliminary analyses revealed two notable findings. First, there was a significant inverse correlation between these two estimates ($r=-.33$, $p=.001$), which is not surprising given the formal complementarity of the focal outcomes in these two questions. Second, participants tended to give higher estimates for the red team ($M=4.80$, $SD=1.11$) than the white team ($M=4.14$, $SD=1.17$), $t(96)=3.50$, $p<.001$. This effect could be attributed to any of a large number of factors that we cannot pinpoint (e.g., the red team was ahead; the color red might inspire more confidence); we will also simply note that this

tendency was not evident in either behavior or predictions (binomial test $ps>.68$; Table 1).

For each participant, we combined the two likelihood estimates into one composite—namely, a difference score reflecting the extent to which his or her estimate for the desired team was larger than for the other team. The average composite estimate was 0.43 ($SD=1.93$), which was greater than 0 and represents a small but significant desirability bias, $t(96)=2.22$, $p=.03$, $d=0.22$. Given the earlier-described main effect in which participants gave higher likelihood estimates for the red versus white team, it is not surprising that the composite estimate for the desirability bias was higher when the desired team was red ($M=1.02$, $SD=1.45$) than white ($M=-0.24$, $SD=2.19$), $t(74.5)=3.30$, $p<.001$, $d=0.68$. More important, the composite estimate for the desirability bias was about the same regardless of whether likelihood judgments were solicited before or after behavior, $t(95)=0.52$, $p=.61$, $d=0.11$.

Correlations, comparison, and mediation analysis

For all analyses reported in this section, we coded all of the relevant variables (desire, likelihood estimates, behavior, and predictions) in a similar direction (higher values meant higher desire for, likelihood estimates favoring, behavior consistent with, and prediction of the red team winning).

We first examined the relations among the three dependent measures. As expected, behavior and predictions were tightly correlated ($r_\phi=.88$, $p<.001$), whereas the correlation between behavior and likelihood judgments ($r=.34$, $p<.001$) was significantly smaller ($p<.001$ for the comparison based on Lee & Preacher, 2013). The correlation between likelihood judgments and predictions was also relatively small ($r=.38$, $p<.001$).

Next, we examined correlations between the desirability variable and the various dependent variables, to allow for comparisons of the effect magnitudes. Redundant with earlier results, the correlations between the desire variable and behavior ($r_\phi=.73$, $p<.001$), predictions ($r_\phi=.69$, $p<.001$), and likelihood judgments ($r=.21$, $p=.04$) were all significant. Comparison of these values reveals that the effect of desire was larger on behavior and predictions than on likelihood judgments (both $ps<.001$; based on Lee & Preacher, 2013).

Finally, we conducted a mediation analysis using the PROCESS macro in SPSS (IBM, Armonk, NY, USA) with 10,000 bootstrap samples and Model 4 (Hayes, 2013, 2014). The desirability manipulation was the independent variable, likelihood judgment was the potential mediator, and behavior was the dependent variable. The 95% confidence interval for the indirect effect did not include zero, $B(SE)=0.39$ (0.32) [0.02, 1.29], which suggests some degree of mediation. However, the confidence interval for the direct effect also did not contain zero, $B(SE)=3.91$ (0.67) [2.59, 5.24]. This suggests that there is a significant effect of desire on behavior, even after accounting for mediation through likelihood judgments.

STUDY 5

The use of the dynamic basketball game as the uncertain outcome in Study 4 was, from an ecological-validity standpoint, an improvement over Studies 2 and 3. However, the basketball outcome was arbitrarily paired with a type of rhyming task. While allowing the study to maintain high internal validity, this arbitrary pairing might be considered peculiar and therefore a limitation. In Study 5, we removed this limitation by telling participants that the outcome of the game influenced whether they would be assigned to give a short speech about “how to win graciously” or “lessons in losing.” Participants were then faced with a choice about whether to read information that would help them prepare for the former topic or the latter topic.

Method*Participants and design*

One hundred thirty undergraduate students at the University of Iowa participated in a 2 (speech target: your team, the opposing team) × 2 (desired team: red, white) × 2 (order: behavior first, likelihood judgments first) between-subjects design.

Procedure, materials, and dependent measures

All procedures and materials were identical to Study 4 except as described here. In place of the rhyming task from Study 4, participants were told that depending on the outcome of the game, they would be required to give an audio-recorded speech about either “how to win graciously” or “lessons in losing.” Half of the participants were told that their speech would be addressed to their own team—so if they won, they would talk about winning, but if their team lost, they would talk about losing. The other half of the participants were told that their speech would be addressed to the opposing team—so if their team won, they would give a speech about losing, but if their team lost, they would talk about winning. This manipulation was necessary in order to counterbalance which speech topic was associated with the desired outcome (i.e., the participant’s own team winning).

For the behavior measure in this study, participants were given an opportunity to read one of two possible articles (prior to knowing the game’s outcome). Participants chose between reading “Things Good Winners Do,” clearly helpful for preparing a speech about winning, and “The Upside of Losing,” helpful for preparing a speech about losing.

The participants did indeed deliver audio-recorded speeches at the end study, but we made this a less stressful experience than what most participants had probably been expecting. Each participant did this in private using a hand-held recording device. The speeches were erased after each session. We added a manipulation check: “Which team do you hope will win?” Two other new measures inquired about the usefulness of the articles participants received and the difficulty of giving the speech.

Results

Results from manipulation check confirmed that participants cared about the outcome and in the expected direction. The response options on that item ranged from 1 = *definitely hope red wins* to 5 = *definitely hope white wins*. Only 11.5% of participants gave the middle response, which was labeled “don’t care.” The mean response was significantly higher among participants for whom the coin flip made the white team “theirs” ($M=4.54, SD=0.60$) rather than the red team ($M=1.73, SD=0.88$), $t(126.81)=21.56, p<.001, d=3.73$. Other preliminary analyses did not reveal noteworthy results regarding trait scales or the additional questions encountered by participants at the end of the session, so those results are not discussed further. Table 2 displays means for the behavior, prediction, and perceived likelihood measures.

Behavior

The effect of desirability on behavior was replicated in this study, although the effect was smaller in magnitude. A significant majority of participants (57.7%) prepared as if the desired outcome would happen ($p=.04$ from a one-tailed binomial test). That is, they chose an article that reflected preparation for the speech that would be required if their team won. This tendency was not significantly qualified by either the team counterbalancing (red vs. white; $p=.41$) or the order counterbalancing (before or after likelihood judgments; $p=.11$). Because of a general overall tendency to prepare for the topic of winning, participants were more likely to prepare for their team winning when they were addressing their team (67.2%) as compared with when they were addressing the opposing team (48.5%), $\chi^2(1)=4.66, p=.03$. This finding, while interesting, should not detract from the main result reflecting an overall tendency to prepare for the desired outcome.

Table 2. Percentages, means, and standard deviations for dependent measures in Study 5

	When red team winning was desired	When white team winning was desired	Overall sample
% Preparing for a red-team win	60.8	46.4	54.6
% Preparing for a white-team win	39.2	53.6	45.4
% Predicting a red-team win	62.2	23.2	45.4
% Predicting a white-team win	37.8	76.8	54.6
Mean likelihood estimate of red-team win (<i>SD</i>)	4.76 (1.03)	4.70 (0.99)	4.73 (1.01)
Mean likelihood estimate of white-team win (<i>SD</i>)	3.93 (1.04)	4.41 (0.99)	4.14 (1.04)

Note: Percentages are in pairs and add to 100. For example, 60.8% of the people who were told the red team was “their team” (i.e., red winning was desired) prepared for the speech associated with the red team winning, while the remainder of those people (39.2%) prepared for the speech associated with the white team winning.

Predictions

Again, as expected, a significant majority of participants (68.5%) predicted that their desired team would win the game ($p < .001$, binomial test). This tendency was not significantly qualified by the team counterbalancing ($p = .08$), the order counterbalancing ($p = .57$), nor whether the speech was supposed to address their own team or the opposing team ($p = .11$).

Likelihood judgments

The average difference between participants' likelihood estimates for the desired and undesired teams was 0.35 ($SD = 1.68$), which was greater than 0 and represents a small but significant desirability bias, $t(129) = 2.35$, $p = .02$, $d = 0.21$. As in Study 4, the composite was higher when the desired team was red ($M = 0.82$, $SD = 1.61$) than white ($M = -0.28$, $SD = 1.55$; $p < .001$). In a borderline effect ($p = .07$), the composite was larger when estimates were solicited before behavior ($M = 0.62$, $SD = 1.59$) rather than after ($M = 0.08$, $SD = 1.73$). There was no significant effect of whether the speech was addressed to the participants' team or the opposing team ($p = .93$).

Correlations, comparison, and mediation analysis

To examine the relations among the three dependent measures, we again coded the relevant variables in the same direction (Study 4). The correlation between behavior and predictions was significant ($r_\phi = .43$, $p < .001$) and was again stronger than the correlation between behavior and likelihood judgments ($r = .18$, $p = .04$; $p = .02$ for the comparison based on Lee & Preacher, 2013). The correlation between likelihood judgments and predictions was again moderate ($r = .47$, $p < .001$).

Next, we compared the effects of desirability by computing correlations between the desirability variable and the various dependent variables. Unlike Study 4, the effect of desire on behavior ($r_\phi = .14$, $p = .10$) was about the same as the effect on likelihood judgments ($r = .17$, $p = .06$). The only significant difference in effects was that the effect on predictions ($r_\phi = .39$, $p < .001$) was greater than the effects on either behavior or likelihood judgment (both $ps < .05$; based on Lee & Preacher, 2013). A similar pattern of correlations was found when we replaced the dichotomous independent variable with the continuous manipulation check variable (i.e., "Which team do you hope will win?"). The correlations of that variable with behavior, likelihood judgments, and predictions were .17, .18, and .40, respectively (all $ps < .05$ with one-tailed tests).

We conducted another mediation analysis using PROCESS Model 4 to determine whether likelihood judgments mediated the relationship between desire and behavior (Hayes, 2013). Perhaps owing to weaker overall effects of desirability in this study relative to Study 4, neither the direct nor indirect effects were statistically significant.

Discussion

The key finding in the previous study was replicated in the current one—participants were more likely to prepare as if the desired outcome rather than undesired outcome was about to happen. Recall that we switched to the speech-preparation task in this study for improved ecological validity of the behavior measure. Anecdotal observations of participants revealed that they were concerned about optimizing their preparation (e.g., they read their chosen article carefully and expressed cogent rationales for their article selections). The desirability manipulation was just one of many possible reasons (unconfounded with desire) for selecting to prepare for one speech topic or another (e.g., knowing less about one topic than the other; feeling that one topic is more important or interesting). Therefore, it is perhaps not surprising that the overall magnitude of the desirability effect on behavior in this study is smaller than in Study 4.

GENERAL DISCUSSION

We began by asking if normatively irrelevant desires would influence how people prepare when faced with an event with an uncertain outcome. We found consistent evidence that people tend to allow the desirability of a key outcome to influence their behavior. They are more likely to act in anticipation of a desired outcome than an undesired one. We showed this over the course of five studies using three distinct paradigms. Study 1 and a follow-up involved hypothetical scenarios of an everyday situation with an uncertain and uncontrollable outcome. Studies 2 and 3 used a computer game to measure immediate behavior in the context of the game. In Study 2, participants prepared more circles on the side of the screen that they hoped an unpredictable slider would go, even though this biased behavior could not influence the final outcome and was a generally inferior strategy. Study 3 replicated this finding using a new desire manipulation (reading a joke vs. solving a difficult math problem) and using grids to communicate likelihood information to participants. For Studies 4 and 5, we used a much different target event—a college basketball game. Both studies, which varied in the preparation task and options presented to participants, revealed that participants' preparation choices depended on outcome desirability.

Our results extend conclusions suggested by studies that have assessed betting behavior within the context of the marked card paradigm (e.g., Irwin & Snodgrass, 1966; Pruitt & Hoge, 1965). As discussed earlier, a potential limitation when drawing conclusions from those studies is that money is used for both the bet and the desirability manipulation, which makes it very obvious to participants that betting on the undesirable outcome would be hedging or betting against the self. It would ensure the gain of a desirable draw would be cut by the loss from the misplaced bet. The present research does not involve bets nor monetary incentives. The hedonic values of outcome desirability and of preparation behavior came in entirely different currencies—with the exception of Study 2, where the desire manipulation and the reward for successful action both involved earning

points. Yet, in all studies, participants still engaged in actions that anticipated their desired outcome.

The direction of our results was not a *fait accompli*. As discussed earlier, various perspectives suggest that desire for an outcome could produce behavior that is, or appears to be, pessimistic. In our studies involving real rather than hypothetical outcomes (Studies 2–5), the key outcomes faced by participants were impending rather than distant. Although previous work has shown that optimism fades (or pessimism grows) as an outcome draws nearer, participants in our studies were still impacted in an optimistic rather than pessimistic direction by the desirability manipulations (Gilovich et al., 1993; Shepperd et al., 1996; Shepperd et al., 2000; Tyler & Rosier, 2009; for a review, see Sweeny & Krizan, 2013). Previous studies on how optimism changes over time have not specifically manipulated desirability and measured its impact on behavior regarding impending outcomes, which make the present work unique. It was also possible that participants in our studies would adopt a pessimistic position because of asymmetric loss functions (the downsides to acting optimistically—e.g., risking disappointment—were worse than the downsides of acting pessimistically; Haselton & Nettle, 2006; Shepperd et al., 1996; Sweeny et al., 2006; Weber, 1994; Weber & Hilton, 1990). The direction of the effects in our studies was not consistent with the direction that this interpretation of asymmetric loss functions would suggest. Finally, although concerns about tempting fate might affect some people on some occasions, such concerns did not—at the level of a main effect—make our participants avoid responding to the desirability manipulations in an optimistic direction (Miller & Taylor, 1995; Risen & Gilovich, 2007, 2008; Swirsky et al., 2011; Van Wolferen et al., 2013). As discussed earlier, previous studies on tempting fate have focused on manipulating whether a protagonist tempts fate rather than on the impact of desirability manipulations.

Despite the very consistent pattern of optimism across our five studies, we do not wish to claim that optimism will always be observed rather than its alternatives of pessimism or a reluctance to tempt fate. There were features of our scenarios and experiments that may have been especially conducive to optimism. For example, in the five studies presented here, the stakes were not very high, and the consequences of behavior that could be construed as “tempting fate” were not very severe. Some research suggests that people are more likely to display optimism for events with minor consequences as opposed to events with severe consequences (Lench & Bench, 2014; Taylor & Shepperd, 1998) and that people are willing to exhibit greater optimism when the consequences of being inaccurate are less severe (Armor & Taylor, 2002). Consequently, if we increased the stakes at play in the present studies, the results might (or might not) shift to show more pessimism or a reluctance to tempt fate.

Likelihood judgments and the biased-guessing account

Three of the experiments (Studies 3–5) included likelihood judgments as dependent measures. The effects of desirability on likelihood judgments were small but significant in all of these studies. The fact that the desirability biases on

likelihood judgments were significant in all three of our studies is somewhat surprising, given that researchers have had only limited success in detecting such effects on likelihood judgments (Bar-Hillel & Budescu, 1995; Bilgin, 2012; Price & Marquez, 2005; Scherer, Windschitl, O’Rourke, & Smith, 2012; Vosgerau, 2010; Windschitl et al., 2013; for a review see Krizan & Windschitl, 2007).

The desirability biases on likelihood judgments also complicate the analysis of whether our results were consistent with an extension of the biased-guessing account. Recall that the account was originally developed to explain documented differences between the sensitivity of likelihood judgments and predictions to desirability biases (Windschitl et al., 2010). The account suggests that when evidence is not decisive, dichotomous outcome predictions can involve hunches or guesswork that are subjectively arbitrary, and this allows for people to insert optimism into their prediction even if they would not exhibit optimism in their likelihood judgments. Likelihood judgments involve scaled responses based on available evidence and may prompt a respondent to think about the probability of potential outcomes relatively deeply, leaving little room for the insertion of hunches or guesses (Windschitl et al., 2010).

Because desirability impacted not only behavior but also likelihood judgments in Studies 3–5, the findings do not fit a simplistic pattern (i.e., null effects of likelihood judgments and significant effects on behavior). Nevertheless, several analyses examining differences between behaviors and likelihood judgments are generally consistent with an extension of the biased-guessing account. First, in Studies 3 and 4, the magnitude of the desirability bias was larger for behaviors than for likelihood judgments. Second, in the two studies in which predictions were measured, behaviors and predictions were more tightly correlated than were behavior and likelihood judgments. Third, in Study 4, the direct effect of desirability on behavior was significant—that is, desirability impacted behavior even after accounting for the biasing effect through likelihood judgment. The application of the biased-guessing account to behavior requires additional work, but the present findings, particularly Studies 3 and 4, provide initial backing for idea that the impact of outcome desirability on behavior is not limited by its impact on likelihood judgment. Although wishful thinking (the causal effect of desirability on likelihood judgment) has been difficult to detect and even considered “elusive” (Bar-Hillel & Budescu, 1995), wishful *behavior* might be more prominent and less elusive.

Future work

In studies on risky choices between gambles, people often reject equal-probability gambles in which the chances of gaining or losing a given amount of money are the same. For these choices, people do not seem to be inclined to behave optimistically—as though uncertainty will be resolved in the desired way. These instances are generally attributed to an asymmetry in the subjective evaluations of gains and losses (the value function in prospect theory; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). If the

gain–loss asymmetry were held constant, might signs of nonindependence between outcome valence and subjective probability be seen? It is possible that people would be inclined to take an equal-probability gamble (a choice that might appear optimistic) except for the fact that the anticipated pain of losing the specific amount is greater than the anticipated enjoyment of winning it. However, recent empirical work by Bilgin (2012) is at odds with this possibility (also Vosgerau, 2010). Bilgin also suggests that outcome valence and subjective probability are not independent, but he argues that losses, rather than gains, “loom more likely.” Our work and Bilgin’s raise some intriguing issues regarding nonindependence between outcome valence and subjective probability. Resolving them is beyond the scope of this paper, but the issues are clearly worthy of further investigation.

Our studies focused on situations in which the evidence for one outcome over the other was indecisive. Based on related empirical work (Krizan & Windschitl, 2007) and theoretical conceptions such as the biased-guessing account (Windschitl et al., 2010), these situations are primary candidates for observing a desirability bias. Finding bias in these situations is important, because life is full of situations in which available evidence is highly indecisive—because evidence is largely missing, the determining event is perceived as random, and/or evidence is well balanced (supports both outcomes about equally). Further research should systematically explore the mitigation of the desirability bias as evidence shifts from being otherwise indecisive to decisive.

More research should also be aimed at the underlying question of why people might tend to take an optimistic stance rather than a pessimistic one in their behavior. The biased-guessing account is primarily about when bias will be observed, and it is shallow in its explanation of why the bias falls in an optimistic rather than pessimistic direction. Here, we will mention two possible explanations regarding the direction of bias in people’s behavior in preparation or anticipation of an outcome. First, perhaps people are simply overgeneralizing from the multitude of situations in which they have control over key outcomes. For many events in their daily lives, people have control over the outcomes and are accustomed to behaving in such a way to make a desired outcome happen. For example, when a student desperately wants an “A” on a test, the student can study intensely in preparation for that test, thereby making the “A” more likely. This logic does not rationally apply to specific cases in which people have no control over the target outcomes (as tested in our studies), but because of simple overgeneralization, people might be prone to behave as if they have control—thereby appearing overoptimistic about the desired outcome.

A second possible explanation refers to the potential averseness of preparing for an undesired outcome. Preparing for such an outcome might be affectively negative because it serves to remind people of something they hope to avoid. For example, making detailed travel arrangements for a work-related trip that one hopes will be unnecessary might be aversive and less likely to get done because the act of making the arrangements brings to mind how unpleasant such travel would be. This idea is related to research on how positive and negative affective reactions to desirable and undesirable

events might underlie optimism and perceived likelihoods (Lench, 2009). In recent work by Lench and Darbor (2014), the negative affect associated with a negative health risk was manipulated via a subliminal priming procedure. Negative affective associations were related to not only decreased perceived risk but also a reduced tendency to take information about reducing the risk. This intriguing finding parallels our findings but with affect as the manipulated variable, and the finding is consistent with the notion that affect may play a key role in the effects of desirability.

While we see promise in these two possible explanations for why people might lean toward optimistic behavior, many other possibilities exist (for additional ideas, see Helweg-Larsen, Sadeghian, & Webb, 2002; Krizan & Windschitl, 2007; Sharot, 2011; Shepperd, Carroll, Grace, & Terry, 2002; Taylor & Brown, 1988; Tyler & Rosier, 2009).

Conclusion

Studying the causal impact of outcome desirability on behavior is most directly accomplished by using an experimental manipulation of desirability for an outcome that is not controllable. Our five studies did just that, across three quite different paradigms. They were consistent in revealing optimistic biases on behavior. Wishful behavior might be relatively common in everyday contexts. Returning to our opening example, it may be that the employee’s desire to participate in the summer traineeship prompts him or her to place a deposit on an apartment whereas a colleague who does not want the position may avoid researching apartments altogether until he or she knows for sure.

APPENDIX A: STUDY 1 FULL SCENARIO— MANIPULATION IN ITALICS

Imagine your house sits in a wooded area with a park behind it. Your city is planning construction of a public pool. One of the potential sites for the pool is in the park behind your house.

[You and your family are super excited about potentially having this pool so close to your house. Your kids, who love swimming and have been begging you to put a pool in your own backyard, would love having such easy access. The proposed plans for the pool look great. The other possible sites for the pool would be much less convenient for your family./ You and your family hate the idea of having this pool so close to your house. You are concerned about the non-stop noise that would emanate from the pool and the heavy foot traffic that it would generate. It would also ruin the beautiful view you currently have of the park and the glare of the lights at night would keep you and your family up. The other possible sites for the pool seem much better for your family.]

The city council already voted to construct the pool in the park behind your house. Therefore, this will be the site of the pool if the results of foundation testing on that site, scheduled in 4 months, are positive. However, there is uncertainty about how the foundation testing will turn out. The engineer in

charge of the test says that it is very hard to tell in advance if the testing will indicate that the site is suitable.

If the pool is built, you will be required by the city to make some landscaping changes to the back end of your property, which would normally cost about \$900. Because of a project that your neighbor is currently doing, you would be able to make the changes at a huge discount. The crew would already be there and you would be able to use some of the same materials, so the project would cost about 70% less than if you brought in a new crew after the results of the foundation testing came back.

APPENDIX B: FULL SCENARIO OF DOG PARK STUDY

Imagine you have recently bought a house with a small yard. Behind it is a grassy area. Your city is planning construction of a dog park.

The city council already voted to put the dog park in the grassy area behind your back yard. However, this plan depends on whether the results of environmental-impact testing are positive. There is uncertainty about how the environmental-impact testing will be positive and allow for the dog park. The person conducting the test says that it is very hard to tell in advance how the dog park might impact the area. They will need to survey the ecosystem of the grassy area, and it will be a few months before the results are in.

If the dog park is built, you will be required by the city to make some landscaping changes to the back end of your property, which would normally cost about \$900. Because of a project that your neighbor is currently doing, you would be able to make the changes at a huge discount. The crew would already be there and you would be able to use some of the same materials, so the project would cost about 60% less than if you brought in a new crew after the results of the environmental-impact testing came back.

ACKNOWLEDGEMENT

This work was aided by support from grants SES 07-20349 and SES 09-61252 to Paul D. Windschitl from the National Science Foundation.

REFERENCES

- Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 334–347). New York, NY: Cambridge University Press.
- Bar-Hillel, M., & Budescu, D. (1995). The elusive wishful thinking effect. *Thinking and Reasoning, 1*, 71–103.
- Bilgin, B. (2012). Losses loom more likely than gains: Propensity to imagine losses increases their subjective probability. *Organizational Behavior and Human Decision Processes, 118*, 203–215.
- Bouts, P., & Van Avermaet, E. (1992). Drawing familiar or unfamiliar cards: Stimulus familiarity, chance orientation, and the illusion of control. *Personality and Social Psychology Bulletin, 18*, 331–335.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*, 319–333.
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*, 306–307.
- Darke, P. R., & Freedman, J. L. (1997). The belief in good luck scale. *Journal of Research in Personality, 31*, 486–511.
- Gilovich, T., Kerr, M., & Medvec, V. H. (1993). Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology, 64*, 552–560.
- Harris, A. J. L., & Hahn, U. (2011). Unrealistic optimism about future events: A cautionary note. *Psychological Review, 118*, 135–154.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*, 47–66.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York, NY: Guilford Press.
- Hayes, A. F. (2014). *PROCESS* (version 2.13). Retrieved from <http://www.afhayes.com/introduction-to-mediation-moderation-and-conditional-process-analysis.html> (accessed on July 8, 2015).
- Helweg-Larsen, M., Sadeghian, P., & Webb, M. S. (2002). The stigma of being pessimistically biased. *Journal of Social and Clinical Psychology, 21*, 92–107.
- Irwin, F. W. (1953). Stated expectations as a function of probability and desirability of outcomes. *Journal of Personality, 21*, 329–335.
- Irwin, F. W., & Snodgrass, J. G. (1966). Effects of independent and dependent outcome values upon bets. *Journal of Experimental Psychology, 41*, 291–297.
- Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods, 6*, 115–134.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47*, 263–291.
- Kogut, T., & Ritov, I. (2011). ‘Protective donation’: When refusing a request for donation increases the sense of vulnerability. *Journal of Experimental Social Psychology, 47*, 1059–1069.
- Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin, 133*, 95–121.
- Krizan, Z., & Windschitl, P. D. (2009). Wishful thinking about the future: Does desire impact optimism? *Social and Personality Psychology Compass, 3*, 227–243.
- Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [computer software]. Available from <http://quantpsy.org> (accessed on July 8, 2015).
- Lench, H. C. (2009). Automatic optimism: The affective basis of judgments about the likelihood of future events. *Journal of Experimental Psychology: General, 138*, 187–200.
- Lench, H. C., & Bench, S. W. (2014). Strength of affective reaction as a signal to think carefully. *Cognition and Emotion, 29*, 220–235.
- Lench, H. C., & Darbor, K. E. (2014). Negative affective reactions reduce perceived likelihood of risk. *Motivation and Emotion, 38*, 569–577.
- Lench, H. C., & Ditto, P. H. (2005). Automatic optimism: Biased use of base rate information for positive and negative events. *Journal of Experimental Social Psychology, 44*, 631–639.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*, 37–44.
- Marks, R. W. (1951). The effect of probability, desirability, and “privilege” on the stated expectations of children. *Journal of Personality, 19*, 332–351.
- Miller, D. T., & Taylor, B. R. (1995). Counterfactual thought, regret, and superstition: How to avoid kicking yourself. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and*

- biases: *The psychology of intuitive judgment* (pp. 334–347). New York, NY: Cambridge University Press.
- Morlock, H. C., & Hertz, J. H. (1964). Effect of the desirability of outcomes on decision making. *Psychological Reports*, *14*, 11–17.
- Price, P. C., & Marquez, C. A. (2005). Wishful thinking in the predictions of a simple repeatable event: Effects on deterministic versus probabilistic predictions. Unpublished manuscript.
- Pruitt, D. G., & Hoge, R. D. (1965). Strength of the relationship between the value of an event and its subjective probability as a function of method of measurement. *Journal of Experimental Psychology*, *69*, 483–489.
- Risen, J. L., & Gilovich, T. (2007). Another look at why people are reluctant to exchange lottery tickets. *Journal of Personality and Social Psychology*, *93*, 12–22.
- Risen, J. L., & Gilovich, T. (2008). Why people are reluctant to tempt fate. *Journal of Personality and Social Psychology*, *95*, 293–307.
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*, 1063–1078.
- Scherer, A. M., Windschitl, P. D., O'Rourke, J. L., & Smith, A. R. (2012). Hoping for more: The influence of outcome desirability on information seeking and predictions about relative quantities. *Cognition*, *125*, 113–117.
- Sharot, T. (2011). *The optimism bias: A tour of the irrationally positive brain*. New York, NY: Pantheon/Random House.
- Shepperd, J. A., Carroll, P., Grace, J., & Terry, M. (2002). Exploring the causes of comparative optimism. *Psychologica Belgica*, *42*, 65–98.
- Shepperd, J. A., Findley-Klein, C., Kwavnick, K. D., Walker, D., & Perez, S. (2000). Bracing for loss. *Journal of Personality and Social Psychology*, *78*, 620–634.
- Shepperd, J. A., Ouellette, J. A., & Fernandez, J. K. (1996). Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback. *Journal of Personality and Social Psychology*, *70*, 844–855.
- Sweeny, K., Carroll, P. J., & Shepperd, J. A. (2006). Is optimism always best? Future outlooks and preparedness. *Current Directions in Psychological Science*, *15*, 302–306.
- Sweeny, K., & Krizan, Z. (2013). Sobering up: A quantitative review of temporal declines in expectations. *Psychological Bulletin*, *139*, 702–724.
- Swirsky, C. L., Fernback, P. M., & Sloman, S. A. (2011). An illusion of control modulates the reluctance to tempt fate. *Judgment and Decision Making*, *6*, 688–696.
- Taylor, K. M., & Shepperd, J. A. (1998). Bracing for the worst: Severity, testing, and feedback timing as moderators of the optimistic bias. *Personality and Social Psychology Bulletin*, *24*, 915–926.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
- Tyler, J. M., & Rosier, J. G. (2009). Examining self-presentation as a motivational explanation for comparative optimism. *Journal of Personality and Social Psychology*, *97*, 716–727.
- Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547–567.
- Van Wolferen, J., Inbar, Y., & Zeelenberg, M. (2013). Magical thinking in predictions of negative events: Evidence for tempting fate but not for a protection effect. *Judgment and Decision Making*, *8*, 45–54.
- Vosgerau, J. (2010). How prevalent is wishful thinking? Misattribution of arousal causes optimism and pessimism in subjective probabilities. *Journal of Experimental Psychology: General*, *139*, 32–48.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.
- Weber, E. U. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin*, *115*, 228–242.
- Weber, E. U., & Hilton, D. J. (1990). Contextual effects in the interpretation of probability words: Perceived base rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 781–789.
- Windschitl, P. D., & Stuart, J. O. (in press). Optimism biases. In Keren, G., & Wu, G. (Eds.), *Blackwell handbook of judgment and decision making: An interdisciplinary perspective*. Malden, MA, USA: Blackwell Publishing Ltd.
- Windschitl, P. D., Scherer, A. M., Smith, A. R., & Rose, J. P. (2013). Why so confident? The influence of outcome desirability on selective exposure and likelihood judgment. *Organizational Behavior and Human Decision Processes*, *120*, 73–86.
- Windschitl, P. D., Smith, A. R., Rose, J. P., & Krizan, Z. (2010). The desirability bias in predictions: Going optimistic without leaving realism. *Organizational Behavior and Human Decision Processes*, *111*, 33–47.

Authors' biographies:

Jillian O'Rourke Stuart is a doctoral candidate in Social Psychology at the University of Iowa. Her research focuses on the social and motivational factors that influence how people think about uncertain events.

Paul D. Windschitl is a Professor of Psychology in the Department of Psychological and Brain Sciences at the University of Iowa. His scholarly interests include judgment, decision making, and social comparison processes.

Andrew R. Smith is an Assistant Professor in the Department of Psychology at Appalachian State University. His research investigates motivational and non-motivational factors influencing people's judgments and decisions.

Aaron Scherer is a postdoctoral research fellow at the Center for Bioethics and Social Sciences in Medicine at the University of Michigan. His research focuses on the factors that bias our beliefs and the consequences of our beliefs, particularly in the domains of health and politics.

Authors' addresses:

Jillian O'Rourke Stuart and Paul D. Windschitl, University of Iowa, USA.

Andrew R. Smith, Appalachian State University, USA.

Aaron Scherer, Center for Bioethics and Social Sciences in Medicine, University of Michigan, USA.