האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

PREDICTING WORLD CUP RESULTS:
DO GOALS SEEM MORE LIKELY WHEN
THEY PAY OFF?

and

WISHFUL THINKING IN PREDICTING WORLD CUP RESULTS:
STILL ELUSIVE

by

MAYA BAR-HILLEL, DAVID V. BUDESCU
and MOTY AMAR

Discussion Paper # 448   March 2007

Following the 2007 publication you can find a 2005 book chapter bearing a different title but covering much the same territory

מרכז לחקר הרציונליות

CENTER FOR THE STUDY
OF RATIONALITY

Predicting World Cup results:
Do goals seem more likely when they pay off?

Maya Bar-Hillel
Hebrew University at Jerusalem, Israel

David V. Budescu
University of Illinois, Champaign, Illinois

AND

Moty Amar
Hebrew University of Jerusalem, Israel
and Ono Academic College, Kiryat Ono, Israel

Bar-Hillel and Budescu (1995) failed to find a desirability bias in probability estimation. The World Cup soccer tournament provided an opportunity to revisit the phenomenon in a context in which desirability biases are notoriously rampant. Participants estimated the probabilities of various teams' winning their upcoming games. They were promised money if one team—randomly designated by the experimenter—won its upcoming game. Participants assigned a higher probability to a victory by their target team than did other participants, whose promised monetary reward was contingent on the victory of its opponent. Prima facie, this seems to be a desirability bias. However, in a follow-up study that made one team salient, without promising monetary rewards, participants also judged their target team to be more likely to win. On grounds of parsimony, we conclude that what appears to be a desirability bias may just be a salience/marking effect, and—although optimism is a robust and ubiquitous human phenomenon—that wishful thinking still remains elusive.

Optimism is one of the most robust and ubiquitous effects in human psychology. Indeed, it may be the hallmark of being well adjusted. The term *depressive realism* refers to the startling possibility that depressed people may not see the world in darker colors than is objectively warranted, but that well-adjusted people see it in rosier colors than is objectively warranted (see, e.g., Alloy & Abramson, 1988). Optimism may take the form of thinking that the world in general is a better place than it really is (e.g., catastrophes are less likely than they really are; future prospects are brighter than they really are, etc.) or that one's personal standing is better than it really is (e.g., "It won't happen to me," etc.).

Evidence for the latter possibility has been obtained in several paradigms. Some studies have compared respondents' estimated chances that desirable personal life events (e.g., high-paying jobs) or undesirable ones (e.g., cancer) would happen to themselves and to their peers (e.g., Weinstein, 1980, 1982). The respondents believed themselves more likely to experience positive life events and less likely to experience negative ones than their peers. Other studies asked for the probability of outcomes of contests (e.g., elections, Babad & Yacobos, 1993; or sporting events, Babad, 1987) in which respondents favored one of the contestants (see also Fischer & Budescu,

1995). Typically, fans inflated the win probabilities of their favored competitors.

A natural explanation for such findings is that the mere desire for a particular outcome is causally sufficient to inflate its judged probability; optimism is a direct consequence of wishful thinking. Bar-Hillel and Budescu (1995) called the causal link between desirability and inflated probability a *desirability effect*. Desirability effects are not necessary conditions for optimism—even of the self-centered kind (although inasmuch as they exist, they contribute to it)—because other mechanisms contribute to optimism. Consider, for example, the privileged knowledge that one has about one's own actions (or one's favorite contestant), or the fact that access to information is often biased and selective. Most people think that their chances of being in a car accident are lower than those of others (McKenna, 1993). Rather than being a reflection of wishful thinking, this may result from any combination of the following: (1) People know more about their own driving skills; (2) people are more aware of their own preventative and cautionary actions; and (3) people pay excessive attention to other drivers' errors.

Without doubting the existence and prevalence of optimism, Bar-Hillel and Budescu (1995) argued that unambiguous demonstration of a desirability effect requires

D. Budescu, dbudescu@uiuc.edu

278

events whose desirability is established through experimental manipulations that are independent of respondents' prior background, knowledge, and preferences. In a series of experiments, they found little evidence that an outcome's desirability—in and of itself—can inflate its judged probability, and they concluded that wishful thinking is an elusive effect.

In a recent literature review, Krizan and Windschitl (2007a) concluded similarly that "despite the prevalence of the idea that desires bias optimism, the empirical evidence regarding this possibility is limited" (p. 95) in studies in which desirability was manipulated experimentally. When the dependent variable was probability judgments regarding aleatory events, their extensive search found only two additional articles (Price & Marquez, 2005; Pruit & Hoge, 1965). These also reported negative findings (however, desirability did affect binary outcome predictions). In naturalistic domains, they found only three additional articles, and even those manipulated desirability somewhat indirectly (Klein, 1999; Krizan & Windschitl, 2007b; Price, 2000). "Overall, there was a small but significant effect of outcome desirability on likelihood judgments . . . and the effects were clearly heterogeneous" (Krizan & Windschitl, 2007a, p. 105). In light of this meager yield, it did not seem redundant to add another study of wishful thinking.

The present article applies Bar-Hillel and Budescu's (1995) paradigm to real-world competitive sports, where wishful thinking seems rampant (see, e.g., Babad & Katz, 1991). Betting is commonplace in sports, with people tending to bet on their favorites and to be overconfident of winning. Our studies were performed during the World Cup soccer games of 2002 and 2006. About 1,300 Israeli students predicted outcomes of some games. Soccer is the most popular team sport in Israel, and games were being broadcast live daily. Thus, even people who did not follow soccer regularly were subjected to frequent updates in the general media. This situation created a felicitous setting for testing the desirability effect yet again, in a context notorious for eliciting wishful thinking.

Results are grouped and reported according to the experimental manipulation rather than chronologically.[1]

THE DESIRABILITY MANIPULATION

Method

Participants. Participants were 800 students, most from The Hebrew University and some from Ono Academic College. Almost half were female. Most were 20–26 years old, with a mean age of 24 years. Students were approached either in classrooms between lectures or in public areas such as cafeterias, and were asked to fill out a short questionnaire regarding the World Cup games, offering an opportunity to win monetary prizes (details below).

Design and Procedure. On June 11 and 12, 2002, participants were handed questionnaires referring to the 8 games to be played on June 13–14, during the last round of the group stage. On June 20, other participants were given questionnaires referring to the 4 games to be played in the quarter-finals on June 21–22. In 2006, between June 6 and June 13, participants were handed questionnaires referring to 5 games from the first round that were scheduled between June 13 and June 21 (the entire round consisted of 48 games that were played June 9–23).

Respondents were asked to estimate for each team the probability of its winning its game. These probabilities had to sum to 100% for teams playing each other (ties could be indicated using 50%). The questionnaires promised respondents several possible rewards. The critical reward did not depend on their performance. Each questionnaire was accompanied by a coupon, stating that 25NIS (then about $4–$5) would be paid to the bearer if the team that was designated on it won. This coupon embodies the desirability manipulation under the assumption that it would make the respondent wish that the designated team would, indeed, win. In order to ensure their attention to the manipulation, respondents were asked to copy the designated team's name onto their questionnaires. This manipulation affected only one of the games judged.

A second payment rewarded accuracy. In 2002, a payment of 25NIS was promised to respondents who correctly predicted the outcome of the largest number of games. In 2006, a payment of 25NIS was promised to participants who correctly predicted the winners of all the games. Predictions were deemed correct if the team that subsequently won had been assigned a probability greater than 50%.

Finally, a much larger prize was promised to the winner of a lottery that was to be held among all of those who won 25NIS by either of the means listed above. In 2002, the prize was 1,000NIS in the first round and 400NIS in the quarter finals. In 2006, it was 500NIS. Thus, respondents could expect payments of 25NIS with a moderate probability and could hope to win the handsome lottery prize.

The questionnaire also queried respondents about their level of interest in the games and their soccer expertise, and asked them to list their favorite team(s). Data from self-reported fans of any one of the teams involved in the manipulations were discarded. Some questionnaires included an additional salience manipulation (described in the following section).

Results and Discussion

Because the study was run on a tight schedule, we managed to manipulate Desirability for just 7 games in 2002 (5 of the 8 games from the last round of the group stage, and 2 of the 4 quarter-finals), and 4 additional games (out of the 5 listed) in 2006. Figure 1 shows the judged probabilities (in percentages) for these 11 games, ordered chronologically (we dropped 26 respondents who failed to note the Desirable team). The abscissa lists the paired teams in each game. The team listed on top is the ultimate winner (Italy–Mexico ended in a tie). The probabilities are those given to the teams on top—when they were named on the accompanying coupon (left bars) versus when their rivals were named (right bars). The number of respondents generating each data point appears next to the team's name.

In 10 of the 11 games, the left bar is higher than the right bar, meaning that when the favorite team was the one whose victory was desired, estimates that it would win were higher than those when the rival team's victory was desired. An ordinal pattern that is this extreme (or more) is statistically significant (with a 12/2,048 probability under the null hypothesis). The overall parametric difference (62% vs. 57%) was also significant $[t(773) = 3.36$, Cohen's $d = 0.24]$, although game by game, the differences were significant for only 2 games (Italy–Mexico; S. Korea–Togo).[2]

These results appear to show a small but systematic Desirability effect, despite the presence of an equally high monetary incentive for accuracy. Before concluding that this is genuine wishful thinking, we must consider an alternative account—that our manipulation simply made the
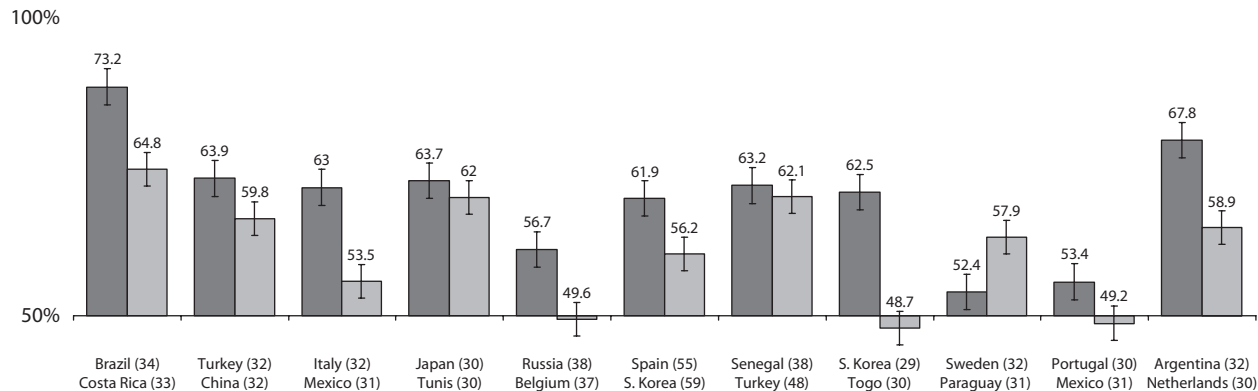
Figure 1. Subjective probabilities that the team listed on top would win, judged by those rewarded if it wins (left bars) and by those rewarded if its opponent wins (right bars). The numbers in parentheses are the sample sizes.

designated team salient, causing it to stand out as being of particular interest. Perhaps it is this "marking" in itself that inflates the probabilities, rather than the fact that the marking happens to have been done by affixing a prize—especially since we forced the respondents' attention upon the rewarded team by asking them to note its name. Such an attentional process is quite distinct from the motivational process implied by wishful thinking.

Related attentional effects have been documented. Barber and Odean (2006)—studying stock-market trading—found that "individual investors are net buyers of attention-grabbing stocks, e.g., stocks in the news" (p. 1). Klar (2002; Giladi & Klar, 2002) found that almost every member of a social group is judged more extreme than the others when singled out as the focal exemplar. Dhar and Simonson (1992) found that an alternative's attractiveness and its choice probability are enhanced when it is made the focus of attention.

In order to test the attention hypothesis, we manipulated salience by simply stating in the questionnaire: *We are particularly interested in team X,* and printing that team's name in boldface. If we were to find a similar effect for the Salience manipulation, parsimony would render desirability superfluous. If drawing attention to a particular team by the simple expedient of expressing interest in it affects

probabilities similarly to drawing attention to a team by naming it on a valuable coupon, then the mere drawing of attention is the critical factor, not the means by which attention is drawn.

THE SALIENCE MANIPULATION

Method

Design. Given the time constraints in 2002, we applied this manipulation to only two games in the quarter-finals of June 21–22 (Spain–S. Korea and Senegal–Turkey). Questionnaires answered by 226 respondents[3] made one team desirable and another team—from the other game—salient, in all four possible combinations.

In 2006, this manipulation was applied alone in 250 additional questionnaires.[4] In order to afford these respondents the same chance to win 25NIS as the Desirability respondents had, we promised them 25NIS if the sum of all goals in the five games would be an odd number.

Results and Discussion

Figure 2—based on 476 respondents—has the same format as Figure 1. Salience enhanced the estimated win probabilities of all teams that were marked by our stated "special interest." This ordinal pattern has a statistically significant probability of 1/64 under the null hypothesis. None of the single-game differences was significant, but over
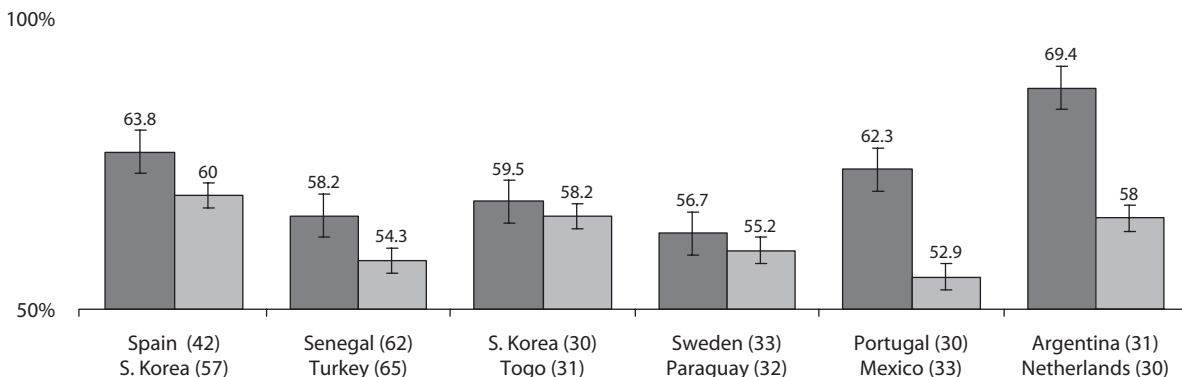

Figure 2. Subjective probabilities that the team listed on top would win, and when its opponent was thus marked (right bars). The numbers in parentheses are the sample sizes.
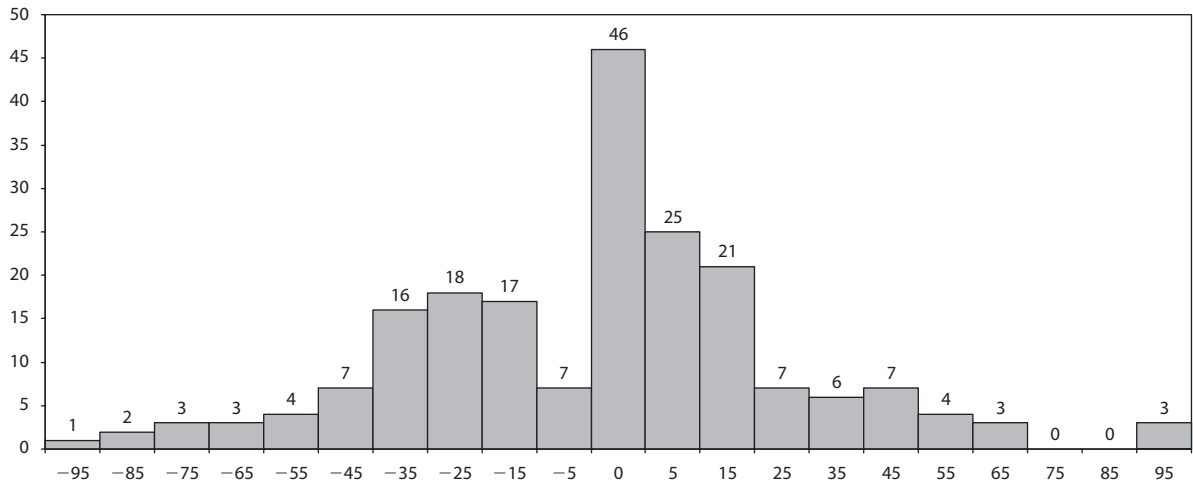
Figure 3. The distribution of the difference scores.

all six games combined, the parametric difference—62% versus 56%—was $t(455) = 3.25$; $d = 0.30$.

The mean magnitude of the Salience effect is essentially identical to that of the Desirability effect—5.2% and 5.4%, respectively. Additionally, among the 12 teams for which we have judgments under both manipulations, Desirability induced higher estimates in 6, and Salience did so in the other 6. Finally, the Desirability and Salience effects did not differ significantly for any team.

The 2002 data allowed a more sensitive comparison of the two effects based on the respondents who were subjected to both. We calculated the difference between the probability they gave to their Desirable team and the probability they gave to their Salient team. If both manipulations have similar effects, then the distribution of the estimates given under the two manipulations should be similar, and the distribution of the differences between them should be symmetric around 0. Figure 3 shows this distribution. Indeed, the mean—0.52%—is not significantly larger than 0: Its median and mode are 0; and the proportion of positive (negative) differences is not significantly different from .5 (binomial test).

There seems to be no justification for positing two separate effects. Our Desirability manipulation obviously marks the targeted team, whereas it is hard to see how the Salience manipulation can affect a team's desirability. Therefore, parsimony compels us to regard the Desirability manipulation as just another way of marking a particular team, thereby making it salient. Conceivably, it had no added function beyond its effect on the respondents' attention.

COMBINING THE MANIPULATIONS

The extent to which the Salience effect resembled the Desirability effect was somewhat surprising, raising the possibility of a ceiling effect (see, e.g., Kunda, 1990). Thus, in 2006, we applied both manipulations to the same team in order to determine if a joint manipulation approaches the ceiling. Note that this cannot speak to whether we are dealing with one attentional and one moti-

vational effect, or just with two attentional effects. Being bitten and scratched hurts more than just being bitten— but so does being bitten twice!

In 261 new questionnaires, the Salience and Desirability manipulations were applied to the same team. Figure 4 shows the results, formatted as those in Figures 1 and 2. The combined effect of both manipulations ($64\% - 53\% = 11\%$) was statistically significant overall [$t(260) = 4.79$; $d = 0.59$], as well as for three individual games (excepting S. Korea–Togo). More importantly, it was more than double the size of the two separate effects ($11\% > 5.2\% + 5.4\%$).

GENERAL DISCUSSION

Our results—small and equal effects of Desirability and of Salience—are robust: They were replicated in practically all of the games considered and were unaffected by the respondent's gender or self-reported interest and expertise in soccer. Had the Desirability results stood alone, we might have concluded that we finally captured the hitherto elusive wishful-thinking effect. The Salience results compel us—for reasons of parsimony—to reject the motivational mechanism: If attention alone can inflate probabilities, then why invoke motivation? In order to show that Desirability can inflate probabilities by itself, we would have had to make some team's victory desirable without singling it out for special attention. We saw no way of doing this in the present paradigm.

How, one might ask, can mere attention to a team inflate the judged probability that it will win? Several related constructs in the literature offer suitable theoretical accounts. Krizan and Windschitl (2007a) listed "possible mechanisms mediating between desires and expectations" (p. 106). Some of them can just as easily mediate between attention and expectations. Thus, the valence priming account posits that "activation of a mental concept spreads more readily to similarly valenced concepts than to differently valenced concepts" (p. 108). The confirmation bias account posits "a specific search strategy . . . for evidence that is consis-
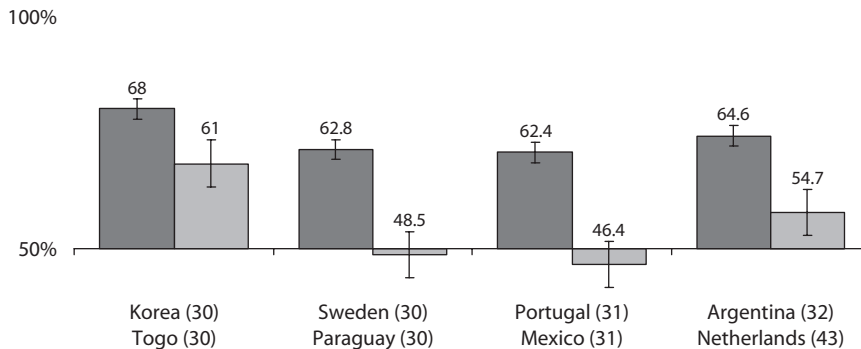
Figure 4. Subjective probabilities that the team listed on top would win when it was both desirable and salient (left bars), and when its opponent was both desirable and salient (right bars). The numbers in parentheses are the sample sizes.

tent with the hypothesis being considered" (pp. 108–109). The focalism account "suggests that . . . people . . . tend to evaluate the evidence . . . relevant to the focal entity while neglecting the evidence relevant to the nonfocal entity" (p.109). These accounts are not mutually exclusive, and the various processes can operate simultaneously. In fact, none of these constructs may suffice by itself to induce bias.

In our judgment task, both the Desirability and the Salience manipulations single out and focus on some team and could cause the respondents to approach the targeted team somewhat differently than any of the other teams (focalism). For example, rather than merely asking themselves in a neutral fashion how likely a team is to win, they might frame the question as: "Can this team win?" Doing so encourages searching for facts about—and features of—the team that could help it win (confirmation bias). They would be considering a win rather than a loss because it is a win that they are asked about, orienting them toward beneficial features even though detrimental features are also relevant (*valence priming*). Respondents would engage in this kind of biased search more with respect to the target team than with respect to the other teams, precisely because it has been targeted (focalism), whether by their presumed desire to see it win (in the Desirability manipulation), or by the experimenters' expressed interest in it (in the Salience manipulation).

These accounts, considered jointly, are consistent with the principle of compatibility that suggests that "selective focusing on features that are compatible with a currently held hypothesis or with the given instructions may be seen to underlie numerous studies reporting . . . confirmatory

biases" (Shafir, 1995, p. 267). They are also consistent with Koehler's (1991) conclusion that "people who explain or imagine a possibility then express greater confidence in the truth of that possibility" (p. 499).

Our results do not contradict magical wishful thinking; they just obviate it. But one could design an experiment in which pure wishful thinking and salience generate conflicting predictions. Suppose, for example, that respondents were rewarded if the target team lost rather than won a game (appropriate background would have to be set for this rather unnatural contingency). The wishful-thinking prediction is that making the team's loss desirable would reduce the judged probability of a win. On the other hand, rewarding the team for a loss makes it no less salient than rewarding it for a win, so it seems that salience predicts that if the probability requested is that for a win, then the probability of a win should rise. A diminished probability of winning would be an argument for the existence of purely motivational wishful thinking.

Having provided an attentional account for our results, the question of why earlier attempts failed to detect it arises. Notably, in rare cases there was evidence for probability inflation. Figure 5 displays results from Study 3 in Bar-Hillel and Budescu (1995). Participants were presented with several hypothetical scenarios involving pairs of contesting parties (e.g., two contracting firms vying for the same bid). Their task was to predict the chances that a given outcome would occur for each competition. Desirability was manipulated by promising a lottery ticket if one of two outcomes occurred. This manipulation biased probability judgments in the expected direction. Study 3—more than any of the others in
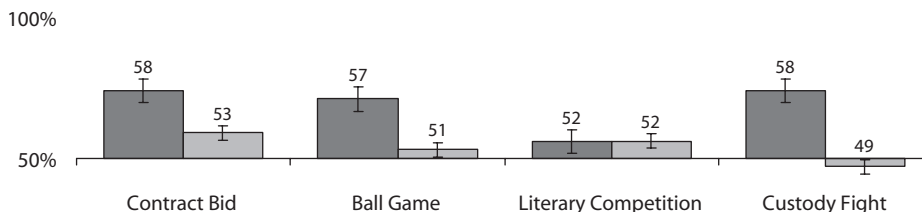

Figure 5. Subjective probabilities that the favorite contestant would win, judged by those rewarded if it wins (left bars), and by those rewarded if its rival contestant wins (right bars).

that article—gave respondents the kind of information that allowed for the combined mental actions described above. Only this study used a scenario paradigm that was rich enough to afford valence priming and manipulated desirability explicitly, affording focalism. Without focalism, contestants are approached symmetrically. Without evidence, valence priming has no grist for its mill. This account can also explain results from other studies. For example, Klein's (1999) Study 1 found that "desirability effects emerged in the evidence condition but not in the no-evidence condition" (Krizan & Windschitl, 2007a, p.104).

The direct causal link "I wish for, therefore I believe in"—which is the essence of "magical" wishful thinking—seems to lack empirical support. The causal link "I focus on, therefore I believe in" has more support. Wishful thinking might, therefore, work indirectly: "I wish for, therefore I focus on, therefore I believe in." This process requires the existence of some evidence on which one can focus when pondering. The widespread belief in the existence of wishful thinking can be explained by the fact that in many real-life situations, people have immediate, direct access to such evidence, and then desires really do cause beliefs to be inflated, via biasing of the evidence.

AUTHOR NOTE

Address correspondence to D. V. Budescu, Department of Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820 (e-mail: dbudescu@uiuc.edu).

REFERENCES

Alloy, L. B., & Abramson, L. Y. (1988). Depressive realism: Four theoretical perspectives. In L. B. Alloy (Ed.), *Cognitive processes in depression* (pp. 223-226). New York: Guilford.

Babad, E. (1987). Wishful thinking and objectivity among sports fans. *Social Behaviour*, **2**, 231-240.

Babad, E., & Katz, Y. (1991). Wishful thinking—against all odds. *Journal of Applied Social Psychology*, **21**, 1921-1938.

Babad, E., & Yacobos, E. (1993). Wish and reality in voters' predictions of election outcomes. *Political Psychology*, **14**, 37-54.

Barber, B. M., & Odean, T. (2006, August). *All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors*. Paper presented at the 2005 EFA Annual Meeting, Moscow. Available from http://ssrn.com/abstract=460660.

Bar-Hillel, M., & Budescu, D.V. (1995). The elusive wishful thinking effect. *Thinking & Reasoning*, **1**, 71-104.

Bar-Hillel, M., Budescu, D. V., & Amar, M. (2008). Wishful thinking in predicting World Cup results: Still elusive. In J. I. Krueger (Ed.), *Rationality and social responsibility: Essays in honor of Robyn Mason Dawes*. New York: Psychology Press.

Dhar, R., & Simonson, I. (1992). The effect of the focus of comparison on consumer preferences. *Journal of Marketing Research*, **29**, 430-440.

Fischer, I., & Budescu, D. V. (1995). Desirability and hindsight biases in predicting results of a multi-party election. In J.P. Caverni, M. Bar-Hillel, F. H. Barron, & H. Jungermann (Eds.), *Contributions to decision making* (Vol. 1, pp. 193-211). Amsterdam: Elsevier, North-Holland.

Giladi. E. E., & Klar, Y. (2002). When standards are wide off the mark: Nonselective superiority and inferiority biases in comparative judgments of objects and concepts. *Journal of Experimental Psychology: General*, **131**, 538-551.

Klar, Y. (2002). Way beyond compare: Nonselective superiority and inferiority biases in judging randomly assigned group members relative to their peers. *Journal of Experimental Social Psychology*, **38**, 331-351.

Klein, W. P. (1999). Justifying optimistic predictions with minimally diagnostic information under conditions of outcome dependency. *Basic & Applied Social Psychology*, **21**, 177-188.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, **110**, 499-519.

Krizan, Z., & Windschitl, P. D. (2007a). The influence of outcome desirability on optimism. *Psychological Bulletin*, **133**, 95-121.

Krizan, Z., & Windschitl, P. D. (2007b). Team allegiance can lead to both optimistic and pessimistic predictions. *Journal of Experimental Social Psychology*, **43**, 327-333.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, **108**, 480-498.

McKenna, F. P. (1993). It won't happen to me: Unrealistic optimism or illusion of control? *British Journal of Psychology*, **84**, 39-50.

Price, P. C. (2000). Wishful thinking in the prediction of competitive outcomes. *Thinking & Reasoning*, **6**, 161-172.

Price, P. C., & Marquez, C. A. (2005). *Wishful thinking in the predictions of a simple repeatable event: Effects of deterministic versus probabilistic predictions*. Manuscript submitted for publication.

Pruit, 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.

Shafir, E. (1995). Compatibility in cognition and decision. In J. R. Busemeyer, R. Hastie, & D. L. Medin, (Eds.), *The psychology of learning and motivation* (Vol. 32, pp. 247-274). San Diego: Academic Press.

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality & Social Psychology*, **39**, 806-820.

Weinstein, N. (1982). Unrealistic optimism about susceptibility to health problems. *Journal of Behavioral Medicine*, **5**, 441-460.

NOTES

1. The results from 2002 were published elsewhere (Bar-Hillel, Budescu, & Amar, 2008).

2. Almost all probabilities given to the top-listed teams were higher than 50%, indicating that the teams that ultimately won were favorites a priori.

3. These respondents were a subset of the 800 that underwent the desirability manipulation, contributing also to Figure 1.

4. Since no team was made desirable in those questionnaires, these respondents are not in Figure 1.

8

WISHFUL THINKING IN PREDICTING
WORLD CUP RESULTS
Still Elusive

Maya Bar-Hillel

The Hebrew University

David V. Budescu

University of Illinois at Urbana-Champaign

Moty Amar

The Hebrew University and Ono Academic College

Bar-Hillel and Budescu (1995) defined the "desirability effect" as "the inflation of the judged probability of desirable events or the diminution of the judged probability of undesirable events" (p. 71). Desirability effects are related to optimism and to wishful thinking, because inasmuch as desirability effects exist, they contribute to both. Evidence supporting these effects was obtained in several types of studies. Some compared respondents' reported chances that desirable personal life events (e.g., getting a high-paying job after graduation) or undesirable ones (e.g., being involved in an accident) would happen to them versus to various others (see e.g., Weinstein, 1980, 1982). Other studies asked participants to judge the probability of outcomes of contests (e.g., elections [Babad & Yakobos, 1993] or sport contests [Babad, 1987]) in which they personally favored one of the contestants (see also Fischer & Budescu, 1995; Granberg & Brent, 1983).

These results, however, are subject to alternative explanations, such as differential information about one's own actions (or one's favorite contestant) versus others' and biased or selective access to information. For example, most people think their chances of being involved in a car accident are lower than others' (McKenna, 1993). Rather than reflecting desirability bias, or unrealistic optimism, this result may be explained by the following facts: (1) people know more about their own driving skills; (2) people are more aware of the preventive and cautionary actions they themselves take; (3) people pay excessive attention to errors made by other drivers; or (4) a combination of all of these factors.

Bar Hillel and Budescu (1995) argued that desirability effects can be demonstrated unambiguously only with respect to events whose desirability is established through experimental manipulations that are independent of the respondents' background, information, and preferences. In an extended series of experimental studies using this paradigm they found little evidence that the desirability of an outcome can, in and of itself, cause its judged probability to loom larger.

In Study 1 Bar-Hillel and Budescu (1995) used aleatory events. Respondents estimated the proportion of cells of a given color in a two-color matrix or of beads of a given color in a jar containing beads in one of two colors. One color was designated as the winning color, but its judged probability of selection did not systematically increase.

In Study 2 the stimuli were four scenarios that concerned, respectively, pairs of firms competing for a contract, finalists in a literary competition, basketball teams playing each other, and parents involved in a child custody battle. One of the competitors in each scenario was made desirable by an appropriate manipulation: The respondent imagined holding stock in one firm (pecuniary); one writer was described as severely handicapped (sympathy); one team was identified as Israel's national team (the respondents were Israeli students); one parent is the same gender as the respondent (in the control condition, the parents' genders were not disclosed in their description). These manipulations had the intended effect on the judged desirability of the target events, but they did not increase the judged probability for the desirable contestant to win. For example, although the respondents wanted the Israeli team to win its game, they did not overestimate the probability of this event.

In Study 3, the same four scenarios were used. Desirability was manipulated by promising respondents a monetary reward if a designated contestant won. This manipulation was the only one that did actually show a significant desirability effect, albeit a mild one. Finally, in Study 4, respondents followed and predicted the value of the Dow Jones index over a four-week period and were rewarded according to whether or not the index changed by 20 points or more (in either direction) over a week. In this study respondents were also rewarded for accurate prediction. Although respondents presumably wished to win the reward, they did not inflate the judged probability of the event that would lead to this outcome. A recent comprehensive literature review (Krizan & Windschitl, 2007) confirmed the scarcity of empirical evidence supporting the desirability bias hypothesis.

Despite a bias against null results in the literature (see, e.g., Sterling, Rosenbaum, & Weinkam, 1995; Hubbard & Armstrong, 1997), these studies were published under a title that told it all: "The elusive wishful thinking effect" (Bar-Hillel and Budescu, 1995).

This chapter reports a field experiment using the same paradigm in the context of real-world competitive sports. Sport is an area in which wishful thinking is notoriously rampant and has been experimentally proven (e.g., Babad & Katz, 1991). Betting in sport is routine and commonplace, with people tending to bet on their own favorites (e.g., the teams they are fans of) and to be overconfident that their bets will succeed (warning against betting on your favorite team is commonplace in guides to sports betting; e.g., www.betinf.com).

A unique opportunity to run this study presented itself in June 2002, when the finals of the World Cup in soccer were being played in Korea and Japan. We conducted an experiment with Israeli participants, involving predictions of the outcomes of World Cup tournament games. Soccer is the most popular team sport in Israel, and most games were being broadcast live in their entirety on a daily basis. Thus, even people who do not follow the sport regularly were subjected to frequent updates on game outcomes in the general news media. This created a felicitous setting for running yet another experiment on the desirability effect but now in a context that elicits involvement quite naturally. It allowed us to test once again whether a paradigm in which desirability is manipulated experimentally would yield desirability bias, this time in a context in which such thinking is commonplace.

STUDY 1

Method

Participants Participants were 329 students at The Hebrew University. About half were female, and almost all were between 20 and 26 years of age, with a mean and median age of 24. They were approached in classrooms at the ends of lectures or in public areas such as cafeterias, libraries, and corridors and asked to fill out a short paper-and-

pencil questionnaire regarding the World Cup games. An opportunity to win considerable monetary prizes (see details below) was promised.

Design and Procedure   Eight games were to be played on June 13 and 14, 2002, as part of the last round of the group stage. On June 11 and 12, 2002, participants were handed a questionnaire that referred to these eight games, in which they were asked to estimate the probability that of each of the 16 teams would win its game. It was pointed out that for teams playing each other, these probabilities had to sum to 100% (predictions of ties could be indicated using 50%).

The questionnaire promised respondents several possible rewards. The reward that was most critical for this study did not depend in any way on the respondent's performance. A coupon accompanying the questionnaire designated one particular team and stated that a sum of 25 NIS (about $5 at that time) would be paid to the bearer of the coupon if the team designated on it won. This coupon embodied the desirability manipulation, under the assumption that it would make the respondent wish that team would win its game. To ensure attention to this team, respondents were asked to write it down on their questionnaires. A second reward of 400 NIS was promised to the winner of a lottery held among all holders of winning coupons (irrespective of the team named on them).

Since these two rewards were not dependent on the respondents' performance, other rewards were promised for the purpose of motivating the respondents to strive for accuracy. Payments of 25 NIS were promised to the respondent who would guess correctly the outcome of the largest number of games. Strictly speaking, the respondents were not asked directly which teams would win, but rather for the probability that a team would win. They were deemed to have predicted the outcome of a game correctly if they assigned the team that subsequently won a probability greater than 50%. Since we expected more than one respondent to tie for this reward, a fourth prize of 400 NIS was to be given by lottery among all those who tied for maximum accuracy. Thus, many respondents could expect payment of 25 NIS, by luck or by competence; they could also hope to win one, or both, of the 400 NIS lottery prizes.

The questionnaire also asked respondents about their overall level of interest in the competition and their level of soccer expertise and asked them to list their favorite team(s).

Results

Because we collected all the data over just two days, our experimental manipulation was limited to five of the eight games played on June

13 and 14, as listed in Figure 8.1. The figure shows the judged probabilities (in percentages) for these games. Each data point was generated by between 31 and 38 respondents (we removed 13 respondents whose favorite team was one of the 10 listed teams; all numbers refer to the remaining respondents).

The team listed on top is the favorite team within each game, namely, the team that, over all respondents, got the higher mean probability for winning. As it turned out, the favorite team was also the winner (Italy and Mexico were tied).

The graph shows the mean probability estimate that the favorite team would win, as given by two experimental groups. The mean estimate of respondents whose coupon indicated that they would win 25 NIS if the top listed team won is shown by diamonds. The mean estimate given by respondents who would be rewarded if the rival team won is shown by squares. By our instructions, rival teams enjoy complementary probabilities. The games on the abscissa are ordered to increase monotonically by the diamonds.

In all five games the diamonds hover above the squares, indicating that the group that presumably wanted the top team to win (because of the desirability manipulation of the coupons) always gave a mean estimate higher than the one given by those who presumably wanted the bottom team to win. This ordinal pattern has a 1/32 probability under the null (chance) hypothesis, so it is statistically significant. Game by
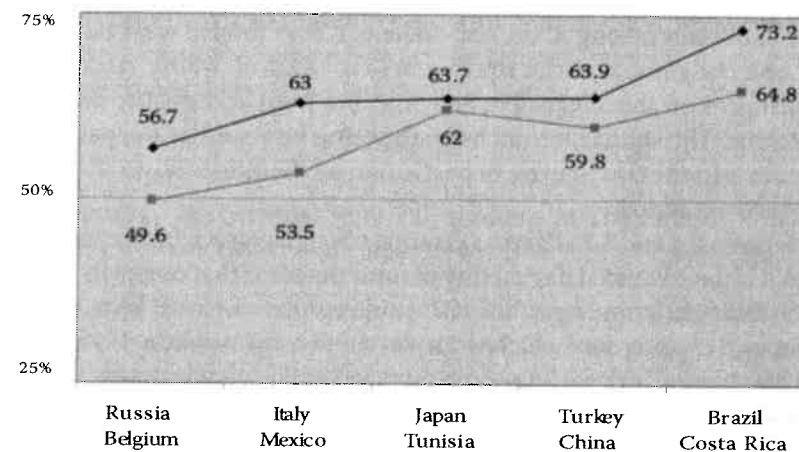

Figure 8.1 Mean probability that the top team listed on the abscissa would win, as judged by those rewarded if it wins (diamonds), and by those rewarded if the rival team wins (squares).

game the parametric differences between the two groups of respondents were statistically significant only for the Italy–Mexico game ($t[61] = 2.35$; $p < .05$; Cohen's $d = .59$). The difference was also significant over all five games (63.9% versus 57.7%; $t[166] = 2.73$; $p < .05$; Cohen's $d = .30$). In other words, designating a team as the one whose win would bestow 25 NIS caused it to seem more likely to win. The effect was not sufficient to reverse the judgments, as 9 of the 10 means are above the 50% line.

Discussion

Figure 8.1 shows a small, but systematic, effect. The outcomes of 1 of 10 teams were endowed at random with value in a controlled between-subject experimental design by making a monetary reward contingent on that outcome. Presumably, this made that outcome desirable. In all five games, an outcome was judged more probable by those respondents who desired it. This effect was found despite the presence of an equally high monetary incentive for accuracy and the absence of any fans of the pertinent teams.

Of all the experiments reported in Bar-Hillel and Budescu (1995), our Study 1 most closely resembles their Study 3. There too the desirability manipulation involved making a monetary reward contingent on one particular outcome out of two possible ones. This also happened to be the only study that yielded what appeared to be a desirability effect. Figure 8.2 shows some of the results of that study in a graph that resembles Figure 8.1 in format (data from Budescu and Bar-Hillel, 1995, Table 7). The diamonds belong to the respondents whose reward was contingent on one outcome, and the squares belong to those whose reward was contingent on the alternative outcome, but both address the diamond outcome. The similar format helps in seeing how similar the pattern of results is in the two figures, in particular, all diamonds hover above the respective squares.

Figures 8.1 and 8.2 suggest a genuine effect. However, before inferring that it is the enhanced desirability of some outcome that causes its judged probability to loom larger, alternative interpretations must be ruled out. One such explanation (which we neglected to consider in the 1995 paper) is that affixing a reward to some outcome makes that outcome salient, singling it out, so to speak, among other outcomes as being of particular interest; and it is this marking in itself that causes the inflated probabilities, rather than the fact that the marking happens to have been done by affixing a prize. Such an attentional process is quite distinct from the motivational process implied by the wishful thinking hypothesis.
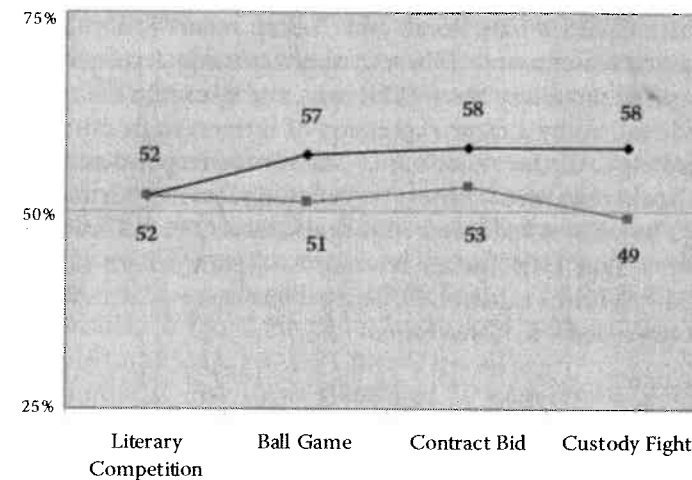
Figure 8.2 Mean probabilities judged by those rewarded if the target event occurs (diamonds), and by those rewarded if its complementary event occurs (squares). Note: Based on data from Bar-Hillel and Budescu, 1995, Table 7.

STUDY 2

To test this possibility, we conducted a second experiment, in which desirability and salience were both manipulated. Naturally, we expected to replicate the effect of the desirability manipulation from Study 1. If there were no similar effect for the salience manipulation, we can attribute the results to a desirability bias with confidence. If, however, we were also to find a similar effect for the salience manipulation, parsimony precludes this interpretation of the results.

Method

Participants   Participants were 227 students at The Hebrew University. They were recruited and run as in Study 1.

Design and Procedure   On June 20, 2002, the participants were handed a questionnaire, which referred to the four games (the quarter-finals) that were to be played on June 21 and 22, 2002. Respondents were asked to estimate the probability of each team winning its game and were promised rewards similar to those in Study 1. In addition to manipulating the desirability of one outcome, another team, playing in another game, was made salient by simply stating, "We are particularly interested in team X" and writing the name Team X in boldface.

Given the time constraints, we applied our manipulations to only two games: Spain versus Korea and Turkey versus Senegal. For each participant a team in one of these games was made desirable by means of a promised monetary reward if it won, and a team in the other game was made salient by a mere expression of interest in its outcome. This generated eight distinct conditions (number of respondents in parentheses): Spain rewarded – Senegal salient (26); Spain rewarded – Turkey salient (29); S. Korea rewarded – Senegal salient (28); S. Korea rewarded – Turkey salient (31); Turkey rewarded – Spain salient (23); Turkey rewarded – S. Korea salient (25); Senegal rewarded – Spain salient (13); Senegal rewarded – S. Korea salient (25).[1]

Results

Figure 8.3 was drawn according to the same guidelines as Figure 8.1. It plots the average probabilities assigned to the events that the favorite teams (i.e., the ones that were predicted to win by a majority of respondents and are listed on top) would win their games.[2] The left panel contrasts the responses of those who stood to win money if the top team won with the responses of those who stood to win money if its opponent won. As in the first study, the former estimates were larger than the latter, albeit not significantly. The right panel contrasts the responses of those who were told that the experimenters were particularly interested in the top team with the responses of those who were told that
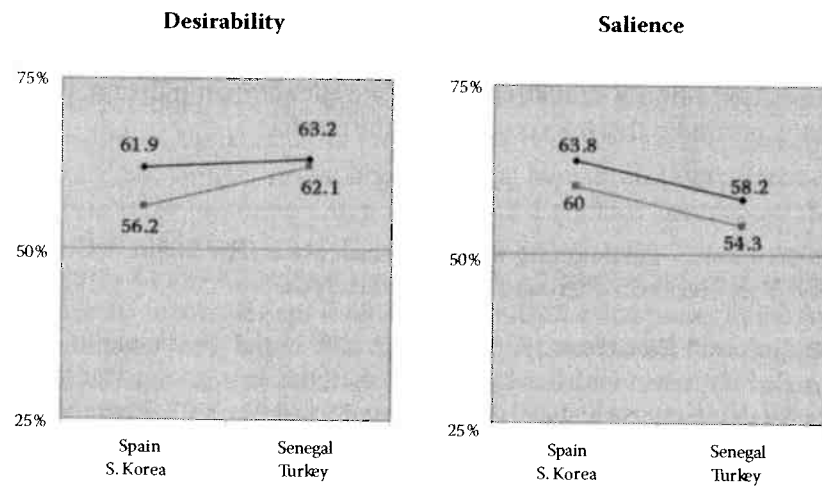
the experimenters were especially interested in its rival team. Again, the former estimates are larger than the latter, but neither difference was statistically significant. As in Study 1, the desirability manipulation is not sufficient to induce an estimation reversal, nor is the salience manipulation.

An eyeball comparison of the two panels shows that the magnitude of the desirability effect is roughly similar to that of the salience effect. We calculated for each team the difference between its mean estimates under the desirability manipulation and the salience manipulation. Two of these differences were negative (–1.9 for Spain, –3.8 for S. Korea), and two were positive (4.9 for Senegal, 7.9 for Turkey), with only the last one significant ($t[111] = 2.13$; $p < .05$, Cohen's $d = .41$).

In an additional analysis we calculated for each respondent the difference between his or her two relevant estimates:

$$d = \text{Prob(team named on coupon)} - \text{Prob(team marked as of special interest)}$$

If both manipulations have essentially the same effect, the mean difference should average about 0. Moreover, the distribution of the differences between the estimates given under the two manipulations should be symmetric. Figure 8.4 shows the distribution of the difference scores. Its mean is 0.52, its median and mode are both 0, and its skewness is 0.03, confirming that the distribution is symmetric around 0.
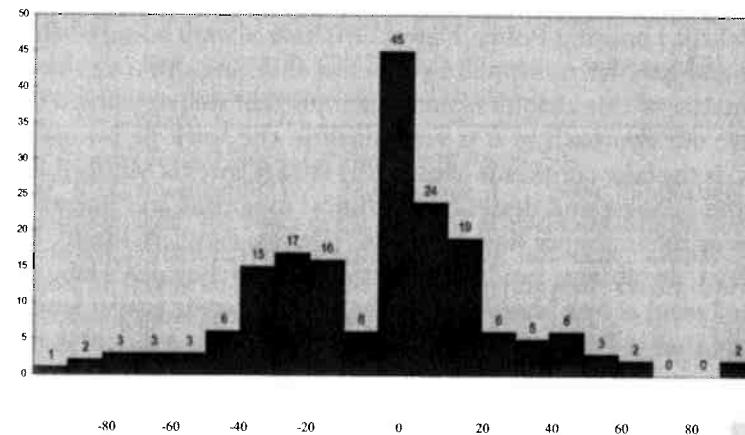
Desirability   **Salience**

Figure 8.3 Mean estimated probabilities of the favorite team winning, as a function of its desirability (left panel) and of its salience (right panel).

Figure 8.4 The distribution of the difference scores.

Discussion

These two studies applied the methodology that Bar-Hillel and Budescu used in 1995 to the present experiments, which were embedded in the real context of the World Cup in soccer. The new results confirm their previous findings and help clarify, at least to some degree, the status of the desirability bias. As in Study 3 in Bar-Hillel and Budescu (1995), we found a small, but robust, effect of the monetary incentives associated with particular outcomes on their judged probability. These results are consistent with the standard finding of desirability bias often observed in natural contexts (Babad & Yakobos, 1993; Babad, 1987; Fischer & Budescu, 1995; Granberg & Brent, 1983; Weinstein, 1980, 1982). However, the results in Figure 8.3 highlight the ambiguity of this interpretation. The "wishful thinking" account implies that this result reflects the operation of a judgmental process that is biased in a systematic fashion by the motivational priming of the desirable outcomes or incentives. Yet, we found practically identical results by merely marking one outcome and making it more salient than the others, absent any motivational manipulation. If salience alone affects probability estimates, is it necessary to invoke a motivational account? Unfortunately, our data cannot answer this question conclusively. Future work should seek to address this issue. Until then, the wishful thinking effect remains elusive.

POSTSCRIPT

We are delighted to have the opportunity to contribute a chapter to this Festschrift honoring Robyn Dawes. We have always admired Robyn's work and have been inspired by his clear thinking. Although the subject matter of this chapter is not on a topic that Robyn worked on, we believe our approach to it is very much in the spirit he brings to his work. Is the false consensus effect really false (Dawes & Mulford, 1996)? Are the richness and flexibility of clinical judgment and judges really superior to "improper linear models" (Dawes, Faust, & Meehl, 1989)? In short, are things, inside and outside the psychologist's lab, always as they seem at first blush? In this chapter (and some earlier work), we did not study whether fans of a soccer team are overconfident that their favorite team will win. They are, and we acknowledge as much, but does this mean they inflate the probability that their favorite team will win *because* they want it to win? We stripped away the context within which sports fans usually operate to ask whether the cognitive mechanism underlying their wishful thinking could possibly be as simple

and direct as the following: When a victory for Team A is exogenously endowed with enhanced desirability, its judged probability increases. Appearances notwithstanding, our answer is no.

REFERENCES

Babad, E. (1987). Wishful thinking and objectivity among sports fans. *Social Behavior, 2*, 231–240.

Babad, E., & Katz, Y. (1991). Wishful thinking—Against all odds. *Journal of Applied Social Psychology, 21*, 1921–1938

Babad, E., & Yakobos, E. (1993). Wish and reality in voters' predictions of election outcomes. *Political Psychology, 14*, 37–54.

Bar-Hillel, M., & Budescu, D. V. (1995). The elusive wishful thinking effect. *Thinking and Reasoning, 1*, 71–104.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*, 1668–1674.

Dawes, R. M. & Mulford, M. (1996). The false consensus effect and overconfidence: Flaws in judgment, or flaws in how we study judgment? *Organizational Behavior and Human Decision Processes, 65*, 201–211.

Fischer, I., & Budescu, D. V. (1995). Desirability and hindsight biases in predicting results of a multi-party election. In J. P. Caverni, M. Bar-Hillel, H. F. Barron, & H. Jungermann (Eds.), *Contributions to decision research I* (pp. 185–203). Amsterdam: Elsevier Science (North Holland).

Granberg, D., & Brent, E. (1983). When prophecy bends: The preference expectation link in the U.S. presidential elections, 1952-1980. *Journal of Personality and Social Psychology, 45*, 477–491.

Hubbard, R., & Armstrong, J. S. (1997). Publication bias against null results. *Psychological Reports, 80*, 337–338.

Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin, 133*, 95–121.

McKenna, F. P. (1993). It won't happen to me: Unrealistic optimism or illusion of control? *British Journal of Psychology, 84*, 39–50.

Sterling, T. D., Rosenbaum, W.L., & Weinkam, J. J. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *The American Statistician, 49*, 108–112.

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*, 806–820.

Weinstein, N. (1982). Unrealistic optimism about susceptibility to health problems. *Journal of Behavioral Medicine, 5*, 441–460.

NOTES

1. Twenty-seven additional respondents contributed data only to the salience results below. We discarded their desirability results because they indicated that they were fans of one of the four teams in the two games we studied.
2. Here, the two favorites happened to lose.