

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/markings 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

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; Pruitt & 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

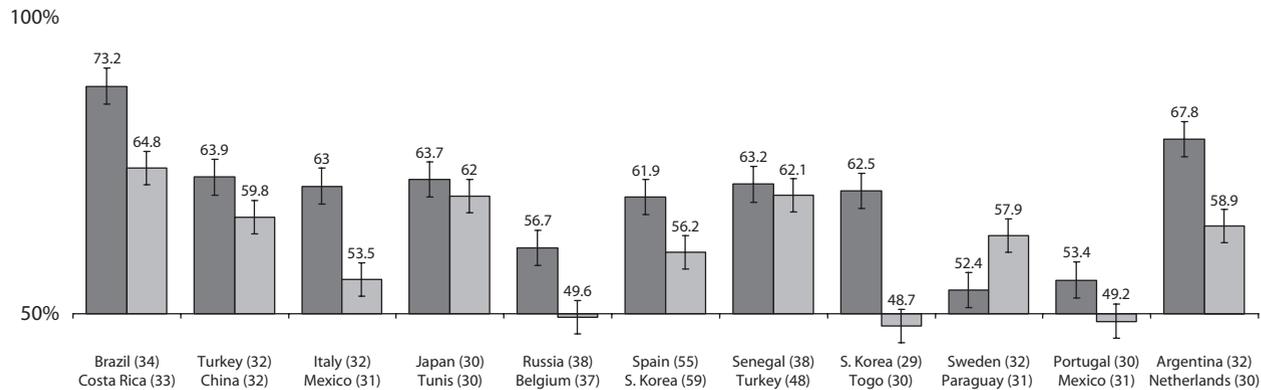


Figure 1. Subjective probabilities that the team listed on top would win, judged by those rewarded if it wins (left bars) and by those rewarded if its opponent wins (right bars). The numbers in parentheses are the sample sizes.

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 respondents³ 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

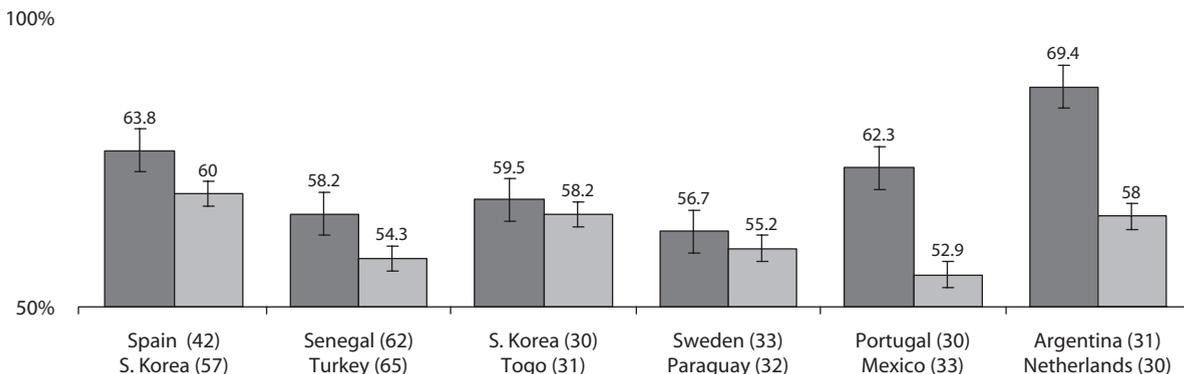


Figure 2. Subjective probabilities that the team listed on top would win, and when its opponent was thus marked (right bars). The numbers in parentheses are the sample sizes.

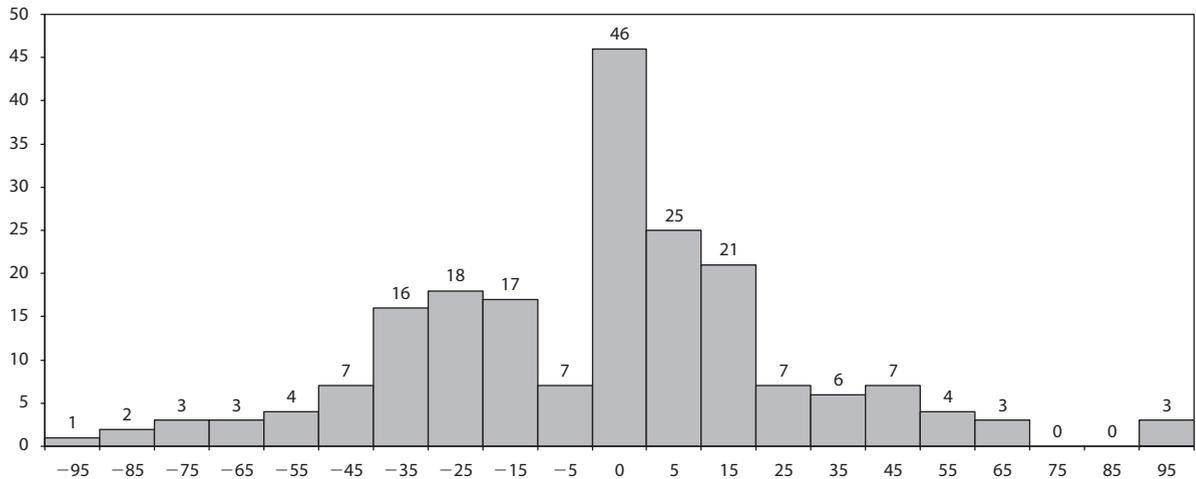


Figure 3. The distribution of the difference scores.

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-

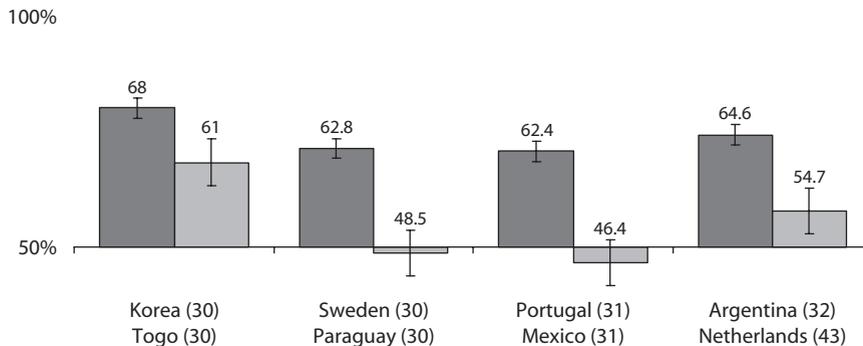


Figure 4. Subjective probabilities that the team listed on top would win when it was both desirable and salient (left bars), and when its opponent was both desirable and salient (right bars). The numbers in parentheses are the sample sizes.

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

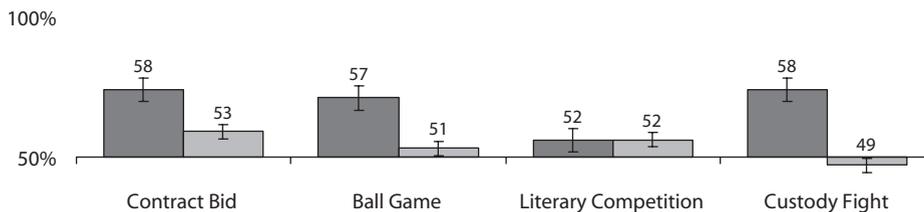


Figure 5. Subjective probabilities that the favorite contestant would win, judged by those rewarded if it wins (left bars), and by those rewarded if its rival contestant wins (right bars).

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.

(Manuscript received March 17, 2007;
revision accepted for publication August 31, 2007.)