Available online at www.sciencedirect.com
SCIENCE DIRECT®
Behaviour Research and Therapy 44 (2006) 875–896
BEHAVIOUR RESEARCH AND THERAPY
www.elsevier.com/locate/brat

Depressive symptoms are associated with unrealistic negative predictions of future life events

Daniel R. Strunk*, Howard Lopez, Robert J. DeRubeis

Department of Psychology, University of Pennsylvania, 3720 Walnut St., Philadelphia, PA 19104-6196, USA

Received 12 November 2004; received in revised form 30 June 2005; accepted 6 July 2005

Abstract

This study examined the relationship between depressive symptoms and bias in the prediction of future life events. Responding to internet announcements, 153 participants varying widely in self-reported depression symptom severity estimated the probability of 40 events occurring over the succeeding 30 days. After the 30-day period, participants reported which events occurred. Optimistic/pessimistic biases were related to level of depressive symptoms. A non-significant optimistic bias characterized participants with low depressive symptoms whereas a significant pessimistic bias characterized participants with high depressive symptoms. Those reporting mild symptoms did not exhibit a systematic pessimistic or optimistic bias. General imprecision in predictions for undesirable events was associated with depressive symptoms. These findings suggest that depression is associated with pessimistic bias rather than accuracy in judgment.

Keywords: Depressive realism; Depression; Cognition; Judgment

Introduction

Beck and his associates propose that depressed people have inaccurate, negative views. Beck asserts that depressed people have cognitive distortions and negative views of the self, world, and future (1967, 1976). Depressed people do report more negative automatic thoughts, more

*Corresponding author. Psychology Department, Vanderbilt University, Wilson Hall, 111 21 St. Avenue South, Nashville, TN 37203, USA. Tel.: +1 615 319 4004; fax: +615 343 8449.
E-mail address: dan.strunk@vanderbilt.edu (D.R. Strunk).

0005-7967/$ - see front matter
doi:10.1016/j.brat.2005.07.001

dysfunctional attitudes, more hopelessness, and a more pessimistic explanatory style than people who are not depressed (Beck, Riskind, Brown, & Steer, 1988; Hill, Oei, & Hill, 1989; Hollon, Kendall, & Lumry, 1986; Peterson & Seligman, 1984). They have also been shown to process negative information more thoroughly and efficiently compared to people who are not depressed (Dozois & Dobson, 2001; Gotlib, Krasnoperova, Yue, & Joormann, 2004). However, Beck's original proposal is that depressed people are not merely more negative than people who are not depressed, but they are inaccurate in their negative views. Testing this claim has proven difficult. A definitive body of empirical work has yet to emerge in this area, despite substantial efforts by several research groups.

Depressive realism

Alloy and Abramson (1979) introduced the depressive realism hypothesis with their report of a series of experiments involving participants with varying levels of depressive symptoms. Their basic experimental procedure involved having college students participate in a series of 40 trials, on each of which participants chose either to press or not press a button, and a green light either did or did not light. The experimenters varied the probabilities of the appearance of the green light after the button was pushed, and not pushed. Alloy and Abramson defined the true degree of control in this situation as the probability of the green light being lit when the participant did press the button minus the probability of the green light being lit when the participant did not press the button. Following these trials, participants were asked to estimate the degree of control their button pressing had over the lighting of the green light. Alloy and Abramson sought to characterize accuracy in judging degree of control over the green light as a function of the level of participants' depressive symptoms.

Although Alloy and Abramson (1979) referred to their dysphoric group as depressed, we use the term dysphoric because participants were drawn from a presumably largely non-depressed sample. While depressed college students would not have been excluded in their studies, the rate of depression in this population along with the relatively low cutoff score they used for the more depressed group suggests that only a very small percentage of their participants were likely to have had moderate to severe depressive symptoms. Alloy and Abramson identified three experimental situations in which a non-dysphoric group exhibited what Alloy and Abramson interpreted as an optimistic bias in their contingency estimates, a bias not evident in their dysphoric group. For example, in one of their experiments, participants had some control (50%) over either winning ten cents per trial or losing 30 cents per trial. The non-dysphoric group underestimated the extent to which they were responsible for their losses, and no bias was evident in the dysphoric group. In all of Alloy and Abramson's experiments, whenever any bias was evident, the non-dysphoric participants' judgments reflected an optimistic tendency that was not present among the dysphoric participants.

These findings have led to a great deal of further consideration and empirical work. And, two versions of the depressive realism hypothesis have now been identified (Ackermann & DeRubeis, 1991). Both versions allow for optimism among people with the lowest levels of depressive symptoms. However, the hypotheses differ in their predictions for people with high levels of depressive symptoms. The first hypothesis, which we shall refer to as the *depressive bias* hypothesis, states that as depressive symptoms increase in severity judgments become more

negatively biased. Under this hypothesis, people with the highest levels of depressive symptoms are expected to exhibit substantial bias and unrealistic pessimism (Beck, 1967, 1976). The second hypothesis, which we shall refer to as the *depressive accuracy* hypothesis, states that as depressive symptoms increase judgments become more and more accurate, such that the judgments of people with the highest levels of depressive symptoms would be the most accurate. These two hypotheses are depicted in Fig. 1. While many psychologists have interpreted existing research to support the depressive accuracy hypothesis, there is reason to doubt the validity of this conclusion.

Three critical limitations of previous work

Because of a number of common limitations, many of the investigations of depressive realism have been inadequate to address whether and to what extent cognition among depressed people is negatively biased, and how any such bias compares to that exhibited by dysphoric or non-depressed people. Several reviewers have highlighted three common limitations of previous research (Ackermann & DeRubeis, 1991; Dobson & Franche, 1989; Haaga & Beck, 1995).

Judgments should be assessed relative to an objective standard not likely to be contaminated by systematic bias

As Ackermann and DeRubeis (1991) have reviewed in detail, one limitation in previous depressive realism work is that many studies have failed to use objective standards in assessing potential bias in judgments. For example, Alloy and Abramson (1988) have suggested that in studies of attributional style dysphoric and depressed "scores for positive and negative events for each dimension did not differ significantly from a score of 0, indicating perfect even-handedness"

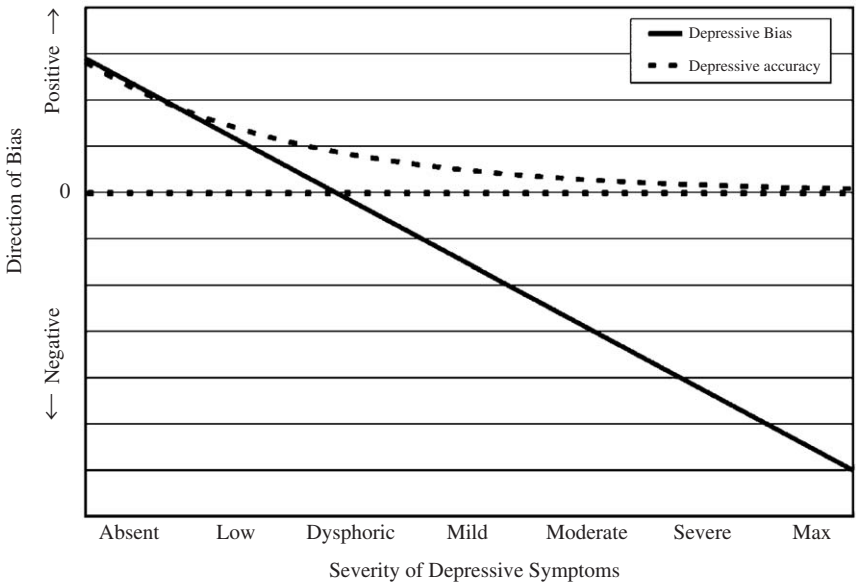

Fig. 1. Two versions of the depressive realism hypothesis: depressive bias and depressive accuracy.

(p. 234). However, there is no evidence that would allow one to relate the *Attributional Style Questionnaire* scores to objective reality (Ackermann & DeRubeis).

Other studies have used measures that were likely to have been influenced by systematic bias. For example, Lewinsohn, Mischel, Chaplin, and Barton (1980) studied judgments of people diagnosed with depression, people with other psychiatric conditions, and people not diagnosed with a psychiatric condition following a small group interaction. They examined group differences in the discrepancies between self and observer ratings of 17 desirable characteristics. While participants were simply asked to rate themselves on the 17 characteristics based on their interactions in the group, the standard used to assess their accuracy was observer ratings. Ackermann and DeRubeis (1991) noted this discrepancy, that observers' ratings have been shown to be critical, and that "it may be the case that the judges' harsh ratings simply happened to coincide with the depressed subjects' low self-ratings" (p. 572). Thus, a number of studies have assessed depressive realism without an objective standard, or with a standard that is likely to have been contaminated by some systematic bias. In the studies in which this limitation was not present, other methodological problems have often clouded the interpretation of results.

Studies should include the full range of depressive symptom severity

A second limitation is that many researchers have assumed that findings based on dysphoric participants will generalize to clinically depressed patients. Researchers have argued that utilizing dysphoric participants to study depression may be inappropriate for a number of reasons (Coyne, 1994; Coyne & Gotlib, 1983). Reviewers of the depressive realism literature have repeatedly lamented the heavy reliance on dysphoric college student samples (Ackermann & DeRubeis, 1991; Dobson & Franche, 1989; Haaga & Beck, 1995). Nevertheless, the judgments of dysphoric people may be of interest in their own right. Researchers might include dysphoric people specifically to investigate the degree of bias among people with mild depressive symptoms rather than as a stand-in for a clinically depressed sample.

Only a few studies have employed samples that include people with a wide range of depressive symptoms. For example, Dobson and Pusch (1995) included depressed people in a non-contingent computerized version of Alloy and Abramson's (1979) procedures. Surprisingly, depressed and non-depressed groups did not differ. Using Alloy and Abramson's paradigm, Carson (2001) found that while depressed participants were consistently more negative via their lower contingency estimates, they were not consistently more accurate. Whereas the non-depressed participants were more accurate in the contingent situation, depressed participants were more accurate in the non-contingent situation. Thus, while the evidence is somewhat mixed, the evidence for the depressive accuracy hypothesis among people who are clinically depressed appears substantially weaker than the evidence generated from studies of dysphoric college students.

Judgments evaluated should be of the type the cognitive model specifies are most relevant

A third common limitation of studies of cognition among depressed and non-depressed groups is the use of artificial laboratory tasks. Reviewers have noted a tendency for artificial laboratory tasks to support depressive realism while more naturalistic, ecologically valid studies have not tended to do so (Ackermann & DeRubeis, 1991; Dobson & Franche, 1989). Beck et al.'s (1979) descriptions of depressed people's judgments have emphasized emotionally salient judgments

related to personal worth. Less ecologically valid studies may not emotionally involve participants sufficiently to test Beck's ideas. Consistent with this idea, Pacini, Muir, and Epstein (1998) have provided some evidence that consequential information is processed differently among dysphoric and non-dysphoric people.

The common methodological limitations of previous work make conclusions about the validity of the depressive accuracy hypothesis premature. The best tests of the cognitive model of depression (and the depressive realism hypothesis) should investigate judgments most relevant to the model. Specifically, these tests should use procedures that examine consequential, emotionally involving judgments relevant to participants' self-worth. Moreover, studies must be conducted in which participants having a wide range of depressive symptom severity and make judgments which are compared to an objective standard to ascertain participants' accuracy.

Design considerations for this study

A study of life event prediction meets these criteria well. In a longitudinal design, participants can report whether events occurred, thereby providing a reasonably unbiased standard by which the accuracy of predictions can be judged. Moreover, the prediction of life events is an everyday, naturally occurring phenomenon. According to the cognitive model of depression, predicting future life events is a theoretically important task. Negative expectations for future events are thought to be a hallmark of depression (Beck, 1967, 1976). Patients' negative predictions for future events are also an important target of cognitive therapy (Beck, Rush, Shaw, & Emery, 1979; Hollon & Garber, 1980). Thus, a study of life event predictions among non-depressed, dysphoric, and depressed participants would provide a test of optimistic or pessimistic bias that utilizes an objective standard, is minimally artificial, and provides a clear test of the cognitive model of depression.

Dunning and Story (1991) examined potential differences in predictions of future life events among groups with different levels of depressive symptoms. In their study, dysphoric participants' probability estimates for life events were found to be less accurate than those of non-depressed participants. Surprisingly, this difference in accuracy was attributable to dysphoric participants being more optimistic than non-depressed participants. This appeared to be due to dysphoric participants failing to anticipate that they would experience fewer desirable events and more undesirable events than non-depressed participants. Although the results reported by Dunning and Story are surprising, the basic methodology when applied to a sample that includes people with more severe depressive symptoms appears appropriate. To our knowledge, no study has addressed the key claims of the cognitive model of depression by investigating bias in life event predictions among people with more severe depressive symptoms.

The purpose of this study is to investigate the relationship of depressive symptoms and pessimistic or optimistic bias while minimizing problems identified with previous research. We investigated this relationship in the context of people estimating the probability of future events in their lives. We endeavored to address four specific issues. First, we hypothesized that participants' tendency to predict positive outcomes (i.e., desirable events occurring and undesirable events not occurring) would be inversely related to the severity of their depressive symptoms. Second, we predicted that the degree to which participants experienced positive outcomes would be inversely related to depressive symptoms. Third, our primary research hypothesis involved the relationship

between pessimistic/optimistic bias and depressive symptoms. This pessimistic/optimistic bias is the bias relative to events that occurred for each participant. We expected that pessimistic bias would characterize the depressed participants, optimistic bias would characterize the non-depressed participants, and a little or no bias would characterize the dysphoric participants. We also planned to explore whether this relationship differed between men and women. Finally, we explored whether depressive symptoms were associated with participants' ability to use their predictions to discriminate between more and less likely events regardless of any optimistic or pessimistic bias they exhibited. That is, we planned to examine the association of depressive symptoms and imprecision in participants' predictions (regardless of their optimistic or pessimistic tendencies).

Methods

Participants

Participants were recruited through internet advertisements on established web sites and postings to newsgroups. The sites and newsgroups were chosen so that some targeted a general audience and others targeted people interested in depression. Therefore, although there were no inclusion criteria for entering the study, this advertisement strategy was employed so that people with a wide range of severity of depressive symptoms would be likely to participate. Although 201 participants completed the initial assessment, 153 participants (76%) completed measures at both assessments and therefore this subset of participants constitute the sample for this study. Ages ranged from 18 to 71 ($M = 35.0$, $SD = 12.5$). The sample was largely composed of women (70% women; 30% men). Participants tended to be well-educated as reflected by their highest degree earned (7% high school diploma; 45% some college; 29% Bachelor's degree; 19% graduate degree).

Based upon simple self-report questions, 28 participants (18% of the sample) reported that they had been recently formally diagnosed with clinical depression. Of these 28 participants, 27 reported that they were currently involved in treatment. No one who reported not being depressed reported being involved in treatment. When asked about being diagnosed with depression at some earlier time in their lives, 25% of those who reported they had recently been diagnosed with depression reported that they had also been formally diagnosed for a previous episode of depression. Of those who reported that they were not currently depressed, 18% reported that they had been formally diagnosed with depression earlier in their lives.

Measures

Depressive symptoms

The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) is a 21-item self-report measure of depressive symptoms. For each item, participants are asked to choose one of four sentences corresponding to the severity of a particular symptom. Items are scored from 0 to 3. The total score is computed by summing the items, and higher scores indicate greater severity of depressive symptoms. The BDI-II appears to be reliable and valid. It has been shown to correlate

highly, $r = .68$, with a commonly used interview assessment of depressive symptoms, the Hamilton Rating Scale for Depression (Beck et al.). Cronbach's alpha for the BDI-II in this study was .94.

Predictions and assessments of life events

Based on previous research, 40 life events were identified as targets of investigation in this study. These events were selected through a series of steps. First, we conducted a review of existing measures of life events to identify potential items. Next, we generated additional items. A subset of 57 items were retained that we judged to be: (a) relevant to the general population; (b) representative of a range of base rates; (c) balanced in the proportion of events that are desirable and undesirable; and (d) representing a range of controllability. Finally, a group of 24 advanced undergraduate students and graduate students evaluated the life events on several dimensions. These dimensions included relevance to the general population, estimated base rate in the general population, desirability of life events (rated on a 1–9 scale), and controllability of life events (rated on a 1–9 scale). Seventeen items were eliminated based upon the judgments of these items provided by the 24 students. A list of the final 40 events (20 desirable and 20 undesirable) selected for use in this study, item means on desirability and controllability ratings and base rates observed for individual events are provided in the appendix.

Participants were asked to make two types of judgments for this list of 40 life events. First, participants were asked to give their estimates of the probability that each of these events would occur to them in the next 30 days. Then, 30 days after these first sets of judgments were made, participants were asked to report whether each event did or did not occur.

Procedure

This study consisted of two parts. Participants were recruited through internet advertisements and offered a chance to win a $200 lottery prize for their participation. All measures were completed via the internet. Participants first read a brief description of the study and provided informed consent. Then participants completed the first part of the study which included completing a demographic information questionnaire, the BDI-II, and predictions of future life events. Twenty-nine days later, participants were sent a reminder e-mail in which they were reminded to complete part two of the study on the following day (day 30). Part two consisted of an internet survey which asked participants to report whether each of the life events of interest had occurred.

Scoring procedures

Several procedures were developed for scoring the data pertaining to participants' predictions of future life events (i.e., their estimates of the probability that each of these events would occur to them in the next 30 days) and their report of whether these events had actually occurred after 30 days.

First, we constructed an index of the extent to which participants predicted positive outcomes. A positive outcome refers to either a desirable event occurring or an undesirable event not occurring. For each event, a number indicating the probability judgment participants assigned to

a positive outcome was computed. For a desirable event, this was simply the probability participants gave for each event occurring. For each undesirable event, the probability of a positive outcome was calculated as one minus the probability participants assigned to that event. The average of all of the probability judgments each participant assigned to each positive outcome was calculated to yield an index of the degree to which participants predicted positive outcomes (called "Predictions").

Second, we constructed a measure of the extent to which positive outcomes occurred after 30 days. Recall that positive outcomes refer to when desirable events occurred or undesirable events did not occur. For each event, a positive outcome was scored with a 1 and a non-positive outcome was scored with a 0. The mean of these event scores will be referred to as Positive Outcomes Experienced (Outcomes). Outcomes scores range from 0 to 1 with higher scores reflecting a greater percentage of positive outcomes experienced.

Third, we created a measure of the extent of pessimistic or optimistic bias in making predictions. As a preliminary step, a score was calculated for each item. For desirable events, the score was the participant's probability judgment minus either 1 if the event did occur after 30 days or 0 if the event did not occur. For undesirable items, this score was calculated as the occurrence/non-occurrence of the event (coded 1 or 0, respectively) minus the probability judgment for the event. The rationale for this scoring can be illustrated with an example. Suppose a participant predicts a 35% chance of a desirable event, and that event does occur. The score of −0.65 (0.35–1.0) indicates pessimism (reflected by the negative sign) at a magnitude of 0.65 on this scale. If the event in this example had not occurred, the score would be 0.35 (0–0.35) indicating optimism (reflected by the positive sign) at a magnitude of 0.35 on this scale. This scoring algorithm results in higher positive scores reflecting greater optimism, and lower negative scores reflecting greater pessimism. Rather than evaluating pessimistic or optimistic bias in making predictions on the basis of a single item, many items were used in this study. These scores computed for each item were averaged to yield the index which will be referred to as "Bias." Bias scores can range from −1 to 1. A score of 0 would indicate that pessimistic and optimistic biases are equal. As scores are further from 0, they indicate greater pessimism (for negative scores) or optimism (for positive scores).

Although the average of these scores calculated for each item may be an intuitively appealing measure of bias, it may be flawed. The flaw is evident when one considers a hypothetical participant response pattern. Suppose some participants gave lower than their honest probability estimates when their true estimate was less than 0.50. And, suppose that these participants also gave higher estimates than their honest probability estimates when their true estimates were greater than 0.50. In short, these hypothetical participants would have given more extreme probability estimates, rather than their best probability estimates. These participants would be rated as more accurate (i.e., less optimistically biased or less pessimistically biased) than their less extreme counterparts, using the scoring method described (Baron, 2000). If this extreme scoring tendency was related to depressive symptoms, this might create a spurious relationship between BDI-II scores and pessimistic or optimistic bias in making predictions. Although participants were not informed of how their results would be scored, and they were instructed to make their best estimates, some participants may still have exhibited this pattern. Thus, although there was no specific reason to suspect that such a response pattern would exist in this study, we considered an alternate scoring method to address this possibility.

Based on the recommendations of Baron (2000), a "proper" scoring rule was employed. Baron explains that when a proper scoring rule is employed, participants would maximize their performance by giving their best estimates. The transformation we applied to these data was the quadratic squaring rule. To apply this rule, the differences computed for each event described above (i.e., the scores calculated in the preliminary step to computing Bias) were squared. The appropriate signs were then applied to these results. More specifically, when either a desirable event occurred or an undesirable event did not occur, the squared value was multiplied by −1. These values were averaged to yield the index which will be referred to as "Bias Squared." Like Bias scores, Bias Squared scores can range from −1 to 1. A score of 0 would indicate that pessimistic and optimistic biases are equal. As scores are further from 0, they indicate greater pessimism (for negative scores) or optimism (for positive scores).

Fourth, an "accuracy" index was calculated using a procedure similar to that used in computing Bias scores. Recall that difference scores were computed for each item in the process of computing Bias. The absolute values of these difference scores for each item were taken. These scores then indicated the magnitude of the discrepancy between participants' predictions and event occurrences for each item. The mean of these 40 scores was then taken. A value of 1 was subtracted from these scores to yield the final accuracy index. These accuracy index scores range from 0 to 1 with higher scores reflecting more accurate predictions.

Results

BDI-II scores ranged from 0 to 47. As shown in Table 1, these scores appeared higher (and more variable) than the Beck Depression Inventory scores reported by Alloy and Abramson (1979) and Dunning and Story (1991). However, this study employed a revised version of the Beck Depression Inventory, the BDI-II, which was not available to investigators working before 1996. Given the differences in the measures used, caution must be exercised in comparing these samples. Research on the psychometric properties of the BDI-II has shown that scores tend to be about 1.54 points higher on the BDI-II than the original BDI (Dozois, Dobson, & Ahnberg, 1998). However, even adjusting for this difference, the scores in this study were significantly higher than those reported by Alloy and Abramson (1979); ($z = 4.20$, $p < .0001$, $d = .47$) and Dunning and Story (1991); ($z = 5.99$, $p < .0001$, $d = .67$). Thus, this study was successful in enlisting participants who, on average, reported more depressive symptoms than previous studies of college students.

For the current study, participants will be divided into three groups on the basis of their BDI-II scores. The range of BDI-II scores for the low, middle, and high-scoring groups are 0–12, 13–19, and 20–63, respectively. The group cutoffs are based on published recommendations for the interpretation of the BDI-II (Dozois et al., 1998; Kendall, Hollon, Beck, Hammen, & Ingram, 1987). The low, middle, and high-scoring groups formed on this basis have been referred to as non-depressed, dysphoric, and depressed, respectively. The cutoffs employed yield scores for non-depressed and dysphoric groups that appear to be roughly comparable to previous work in college students (see Table 1).

The high-scoring group in the current study was not diagnosed with MDD as part of this study. In an effort to make this point clear, we will refer to these groups as low, middle, and high-scoring groups. Despite not being formally diagnosed, the high-scoring group in this study reported

Table 1
Beck Depression Inventory scores for the current study compared to Alloy and Abramson (1979) and Dunning and Story (1991)

	Overall			Groups by symptom severity								
				Low (Non-depressed)			Middle (Dysphoric)			High (Depressed)		
	M	SD	*N*	*M*	SD	*N*	*M*	SD	*N*	*M*	SD	*N*
Current study	14.3	11.2	153	5.9	3.5	83	16.7	2.0	31	30.5	6.9	39
Dunning and Story (1991)	7.2	3.9	423	5.1	3.0	330	15.1	5.0	89	—	—	—
Alloy and Abramson (1979)	10.3	4.4	288	5.3	3.0	144	15.2	6.3	144	—	—	—

Note: For data not from the current study, means reported are weighted averages of subgroup means authors reported and standard deviations are pooled standard deviations of subgroup standard deviations authors reported. Mean Beck Depression Inventory scores for Dunning and Story (1991) and Alloy and Abramson (1979) have been increased by 1.54 points to facilitate comparisons between the two versions of the inventory (Dozois et al., 1998).
For Dunning and Story (1991) and Alloy and Abramson (1979), groups were formed using an original BDI score of 9 or greater as the cutoff point. Therefore it is possible that some participants who scored very high on the BDI were included in the dysphoric group in these studies.

depressive symptoms at least as severe as those reported by patients in clinical trials of treatments of depression. The BDI scores of participants in the Treatment of Depression Collaborative Research Program (TDCRP) had a mean of 26.8 and a standard deviation of 7.9 (Elkin et al., 1989). After adjusting for the mean difference between the older and newer Beck Depression Inventories, the BDI-II scores in the present study are significantly higher than the BDI scores of the depressed group in the TDCRP ($z = 1.70$, $p < .05$, $d = .29$). Thus, using BDI scores, a substantial portion of the participants in the current study reported depressive symptoms at or above the levels reported by depressed patients.

Using the cutoffs described above, approximately 54% of the sample was categorized as low-scoring (non-depressed), 20% was categorized as middle-scoring (dysphoric), and 25% was categorized as high-scoring (depressed). Thus, each category was reasonably well represented. This will facilitate a more powerful test than if the some categories had been poorly represented. This may have been a problem in Dunning and Story's (1991) study in which 78% of their sample was categorized as non-depressed, with the remainder being dysphoric.

Predictions of future events and base rates of events studied

Before proceeding to testing specific hypotheses, some description of the prediction and event occurrence data is warranted. The average probability given for future events was .35 with a standard deviation of .36. On average, at the 1-month follow-up, participants reported that 32% of the 40 events occurred (SD = .47).

It is also worth considering these data broken down by the desirability of events. Participants gave higher probability estimates for desirable events ($M = .47$, SD = .37) than undesirable events ($M = .22$, SD = .30, t (1, 152) = 15.35, $p < .0001$). Closely paralleling this difference,

participants reported that a higher percentage of desirable events ($M = .45$, SD = .50) occurred relative to undesirable events ($M = .18$, SD = .38, t (1, 152) = 16.90, $p < .0001$). The latter finding shows that the base rates of desirable and undesirable events targeted in this study differed. Given this difference, the difference noted in the probability estimates given for desirable and undesirable events may not be surprising. However, this is a different question than asking whether Bias, which takes the events that each participant experienced into account, differed between desirable and undesirable events. The Bias for desirable events ($M = .02$, SD = .45) was significantly higher (indicating greater optimism) than the Bias for undesirable events ($M = -.04$, SD = .34, t (1, 152) = 2.97, $p < .01$, $d = .15$). Nonetheless, because Bias is so integral to testing primary study hypotheses, this finding suggests that it may be important to examine findings to ensure that there are not other differences between desirable and undesirable events. Specifically, the relationship between Bias and BDI-II scores will be examined separately for desirable and undesirable events. To do so, the average of the scores calculated as the preliminary step to calculating bias (described previously in "Scoring Procedures") was taken separately for desirable and undesirable events.

Analysis of the internal consistency of scales constructed provides further evidence to suggest the potential importance of examining desirable and undesirable events separately. First, Cronbach's alpha was computed for the 40 scores that indicated the extent of the positive predictions that were used in calculating Predictions. Cronbach's alpha for Predictions was .79. Cronbach's alphas for the two separate Predictions scales for desirable and undesirable events were both high (.89 and .86, respectively). Second, the internal consistency score from Kuder–Richardson formula 20 (KR-20) was computed for the 40 scores indicating the extent of positive outcomes in the occurrence of events. These were the scores used in computing Outcomes. KR-20 for Outcomes was .60. KR-20s for Outcomes for desirable and undesirable events were .73 and .49, respectively. Finally, Cronbach's alpha was computed for the 40 scores indicating the extent of pessimistic or optimistic bias in predicting events. Cronbach's alpha for Bias was .49. Cronbach's alphas for Bias for desirable and undesirable events were .74 and .66, respectively. Generally, Cronbach's alpha is expected to increase with greater numbers of test items. The lower Cronbach's alpha for the overall Bias appears to be attributable to a negative correlation between Bias for desirable events and Bias for undesirable events ($r = -.41$, $p < .0001$). This negative relationship and the fact that Cronbach's alpha for the overall Bias is lower than the Cronbach's alpha for both Bias for desirable events and Bias for undesirable events suggest that it may be important to examine Bias for desirable event and Bias for undesirable events scales separately.

The mean of Bias across all participants ($M = -.01$, SD = .08) was only slightly lower than zero, but it indicated a significant tendency towards pessimism in this sample ($t = -2.48$, $p < .05$). This should be interpreted in the context of a study that actively recruited people with greater than average levels of depressive symptoms.

Depressive symptoms predicting life event predictions and occurrence of events

Although it is not the most central research question, we first examined the relationship between Predictions and BDI-II scores. This relationship between predicting good outcomes (regardless of event occurrence) and BDI-II was significant ($r = -.59$, $p < .0001$). This is a large

effect, indicating that people with greater depressive symptoms were substantially less likely to predict positive outcomes.

Next, we examined whether participants with higher BDI-II scores tended to report fewer positive outcomes at the end of the 30-day period. Outcomes was significantly associated with BDI-II scores ($r = -.30$, $p = .0002$). This is a medium-sized effect. Thus, people who scored higher on the BDI-II tended to report having experienced a combination of more undesirable events and fewer desirable events.

To address our primary research hypothesis, we examined the correlation between Bias and BDI-II scores. Consistent with our hypothesis, Bias was moderately and significantly related to BDI-II scores ($r = -.35$, $p < .0001$). This is a medium effect. Greater pessimism/less optimism was associated with more severe depressive symptoms. The scatter plot of this relationship is depicted in Fig. 2. Deletion of possible outliers in this analysis did not substantially influence the correlation coefficient obtained. These variables were then examined in a linear regression analysis. The intercept of that linear model (13.7) is the BDI-II score associated with the absence of any optimistic or pessimistic bias. This BDI-II score of 13.7 is quite close to the mean BDI-II score in the middle-scoring (dysphoric) group. The intercept of a linear model in which the independent variable and dependent variable were reversed was .03. This indicates that a participant with a score of 0 on the BDI-II would be expected to have an optimistic bias of .03. Recall we also calculated a version of Bias called Bias Squared. The relationship between Bias Squared and BDI-II scores was very similar ($r = -.38$, $p < .0001$).

A series of *t*-tests was conducted to examine, in each of the three separate groups (i.e., low-scoring, middle-scoring, and high-scoring on the BDI-II), whether the mean Bias differed from

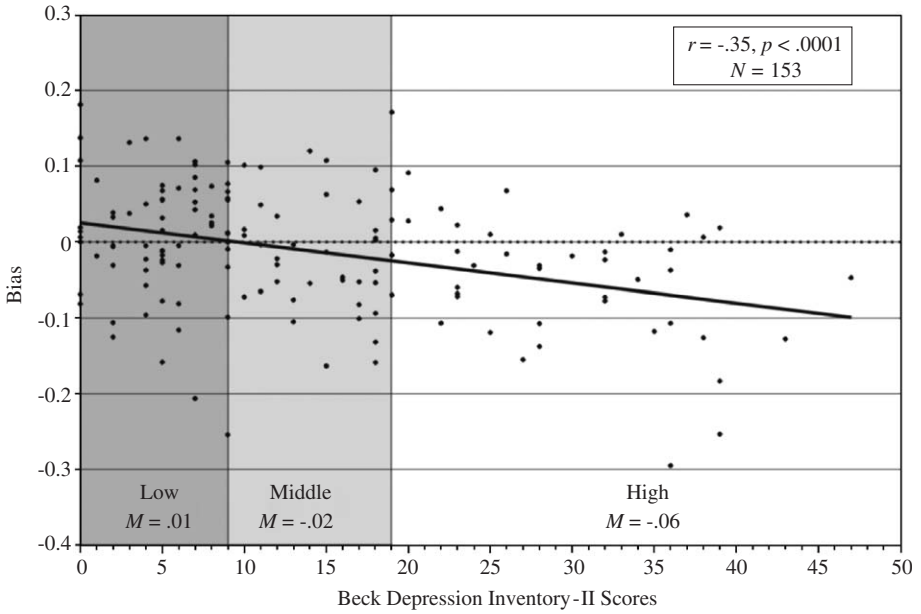

Fig. 2. Pessimistic/optimistic bias and depressive symptoms.

zero. The high-scoring group was significantly pessimistic ($M = -.06$, SD = .08, $t = -4.27$, $p < .0001$). The middle-scoring group exhibited neither a significant optimistic or pessimistic tendency ($M = -.02$, SD = .08, $t = -1.05$, ns). Finally, the low-scoring group showed an optimistic tendency that was not significantly different from zero ($M = .01$, SD = .08, $t = 1.02$, ns).

To provide a more direct test of the depressive accuracy hypothesis, a quadratic term was tested in a regression analysis of Bias as a predictor of BDI-II scores. As Fig. 1 illustrates, the depressive bias hypothesis predicts a linear relationship between Bias and BDI-II scores, whereas the depressive accuracy hypothesis predicts a curvilinear relationship. A quadratic term was examined. This term was nearly zero ($\beta = .01$) and not significant ($p = .85$). Therefore, these data fail to support the depressive accuracy hypothesis.

To facilitate comparison with the research that has utilized a more restricted range of depression severity, the distribution of BDI-II scores in the present study was truncated to approximate the mean reported by Dunning and Story (1991; adjusted for the mean difference expected between the two versions of the Beck Depression Inventory). A subset of participants was selected for inclusion in this analysis. A normal distribution was assumed. Nevertheless, within these parameters, a very large number of subsets was possible. We chose to assign participants z-scores based upon the mean and standard deviation reported by Dunning and Story. For each one standard deviation unit of the distribution, participants were randomly selected so that each standard deviation unit of the distribution was represented in proportion to the normal distribution. For example, 34 participants were selected who had z-scores between 0 and 1, whereas 14 participants were selected who had a z-score between 1 and 2. Using this strategy, the sample was reduced to 102. The resulting BDI-II scores ($M = 8.4$, SD = 5.5) were more comparable to those of Dunning and Story. The exclusion of higher scoring participants led to a reduction in the correlation of Bias and BDI-II scores from $r = -.35$ in the full sample to $r = -.15$ in the reduced sample, which was not significant (see Fig. 3). Thus, when participants were removed so as to mimic the distribution of BDI scores in Dunning and Story, a positive result was not obtained, whereas use of the full sample led to a significant relation between BDI score and bias.

Possible alternative explanations to relationship between bias and BDI-II

The relationship between BDI-II scores and Bias was examined separately for desirable and undesirable events, given the lower internal consistency for overall Bias relative to the internal consistencies of Bias for desirable events and Bias for undesirable events. The relationship between Bias and BDI-II scores appeared to be very similar for desirable and undesirable events ($r = -.21$, $p = .01$ and $r = -.19$, $p = .02$, respectively). Each component of the Bias total score appears to have contributed approximately equally to the relationship between overall Bias and BDI-II scores. The relationships did not differ significantly ($z = 1.68$, ns). Thus, this pattern of results suggests that the relationship between BDI-II scores and Bias was not driven by only desirable (or undesirable) events.

We conducted a regression analysis to examine whether the relationship between Bias and BDI differed between men and women. Previous research has shown that men and women differ in depressive symptom severity and the ways they cope with depressive symptoms

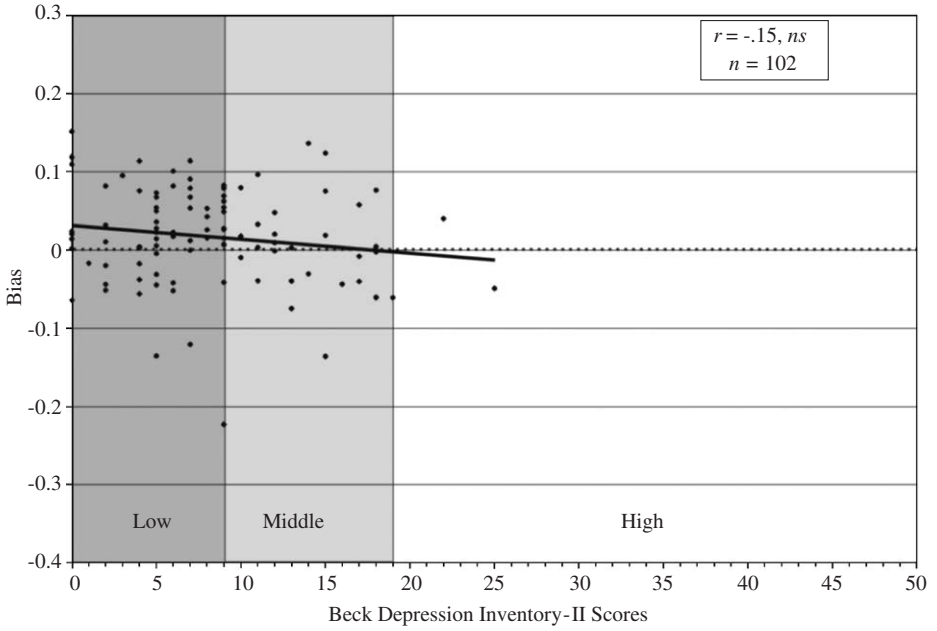

Fig. 3. Pessimistic/optimistic bias and depressive symptoms in reduced sample.

(Nolen-Hoeksema, 1987). Bias and gender were centered prior to examining these variables and their interaction as predictors of BDI-II scores. The interaction of Bias and gender was significant ($\beta = .15$, $p < .05$) indicating that the relationship between Bias and BDI-II differed significantly for men and women. To illustrate the magnitude of this effect, we examined the correlations between Bias and BDI-II for men and women separately. The correlation for women ($r = -.39$) was larger than that for men ($r = -.14$).

Accuracy in predictions

An analysis was conducted to examine whether depressive symptoms were related to accuracy in predictions. Recall that "accuracy" is defined as the magnitude of the discrepancy between the predictions participants made and their report of which events occurred. BDI-II scores were related to the accuracy of predictions ($r = -.21$, $p = .01$). Subsequent analyses revealed that this relationship was present for undesirable ($r = -.32$, $p < .001$), but not for desirable events ($r = -.01$, ns).

Bias was then entered as a covariate to ensure that any relationship between accuracy and BDI-II scores was not merely driven by Bias scores. Thus, the analysis of the relationship between BDI-II scores and accuracy addressed the issue of whether depressive symptoms are associated with poorer, more inaccurate predictions, above and beyond any inaccuracy that may have been associated with having a pessimistic or optimistic bias. After controlling for Bias, BDI-II scores were still marginally related to the accuracy of predictions ($r = -.14$, $p = .06$). Controlling for Bias, BDI-II scores were not related to accuracy in predicting desirable events ($r = -.01$, ns), but

were related to accuracy in predicting undesirable events ($r = -.29$, $p < .001$). Thus, depressive symptoms were negatively associated with accuracy in predicting undesirable events.

Discussion

This study found support for the association between pessimistic or optimistic bias and level of depressive symptoms. Participants who were high-scoring on depressive symptoms exhibited a pessimistic bias. Participants who were low-scoring on depressive symptoms exhibited an optimistic bias (though not significantly so). Middle-scoring participants exhibited the least pessimistic or optimistic biases.

Methodological strengths

This study examined the relationship between severity of depressive symptoms and bias in predicting life events over a 30-day period. The study had three primary methodological strengths. First, participants in this study reported a wide range of depressive symptoms. Compared to studies utilizing samples of presumably largely non-depressed college students, this study appeared to obtain participants with more severe depressive symptoms. In fact, the high-scoring group of participants in this study reported BDI-II scores more severe than those reported in the Treatment for Depression Collaborative Research Program (Elkin et al., 1989). The inclusion of participants with such a wide range of depressive symptoms allows for greater power in detecting any relationship involving depressive symptoms, and it allows the results to be generalized to a wider range of depression severity.

Second, the type of judgments investigated in this study is critically important in the cognitive model of depression. These judgments are predictions of one's personal future. More specifically, participants are asked to predict events that are desirable and undesirable. Because the events used were not neutral events, the prediction task was likely to involve participants emotionally. In fact, as noted previously, much of the cognitive theory of depression deals specifically with expectations for future events (Beck et al., 1979; Hollon & Garber, 1980).

Third, participants provided a reasonably unbiased measure of the accuracy of their predictions by reporting if each of the events of interest in this study had occurred 30 days after they made their predictions. Insofar as the reports given by participants were veridical, this study had an objective measure of reality. Taken together, this study had a unique combination of important methodological strengths that make it an excellent opportunity for addressing the primary research question: Are depressive symptoms associated with unrealistic pessimism in predicting future events?

Depressive symptoms, predictions of positive outcomes, and positive outcomes experienced

Consistent with a wealth of previous research demonstrating that depressive symptoms are associated with negative thinking patterns, depressive symptoms reported by participants in this study were strongly related to lower predictions of positive outcomes (i.e., desirable events occurring and undesirable events not occurring). Thus, depressive symptoms were related to

negativity in predictions. Of course, this does not address the primary question of whether these predictions were truly more pessimistic, compared to an objective standard.

Also, consistent with previous research (Farmer & McGuffin, 2003; Hammen, 1991), greater depressive symptoms were associated with less frequent occurrence of desirable events and more frequent occurrence of undesirable events. As Hammen (1991) has pointed out, the more frequent occurrence of undesirable events and the less frequent occurrence of desirable events may be due to a variety of contributing factors. The more negative life events experienced may contribute to participants' depressive symptoms. And, depressive symptoms appear to increase people's chances of experiencing more negative life events (Coyne, 1976; Hammen, 1991). It is important to recognize these sources of low positive outcomes in life events. However, in understanding this study, it is important to note that participants were free to use any information about themselves or their world to make their predictions. Therefore, the high-scoring participants predicted a negative future that was not borne out by their experience.

Depressive symptoms and pessimistic or optimistic bias in predictions

This pattern of results allows a number of previous findings to be integrated. First, the finding of optimistic bias among those with low levels of depressive symptoms (as is seen in the general population) is consistent with the research reviewed by Taylor and Brown (1989) showing that, in general, people tend to show a host of optimistic biases.

Second, these findings are also consistent with the majority of depressive realism research which has relied largely on non-depressed college student samples (Ackermann & DeRubeis, 1991; Dobson & Franche, 1989). In these studies, the college students with higher levels of depressive symptoms (comparable to the middle-scoring, or dysphoric, group in this study) have tended to show an absence of a pessimistic bias compared to their low-scoring, non-depressed counterparts. Within this study, it is also interesting to note that our effort to truncate the mean of the distribution of scores on the BDI-II led to a failure to find a relationship between BDI-II scores and pessimistic bias. Thus, studies that have lacked participants with more severe depressive symptoms have been poorly powered to detect the relationship between depressive symptoms and pessimistic or optimistic bias. Moreover, by not including participants to adequately represent the full range of depressive symptoms, these researchers could not properly estimate the extent of pessimistic or optimistic bias associated with high levels of depressive symptoms. The current study was able to accomplish this goal.

Finally, the current findings can be compared to findings on the extent of pessimistic or optimistic bias in research that has included participants with high levels of depressive symptoms. For example, recall that two studies (Carson, 2001; Dobson & Pusch, 1995) used the Alloy and Abramson (1979) paradigm in a sample of depressed people. Dobson and Pusch (1995) failed to find any differences in contingency judgments between depressed and never depressed participants. Carson (2001) found that depressed participants consistently gave lower estimates of contingency. These predictions tended to be accurate in a non-contingent situation and inaccurate in a contingent situation. Given this latter finding demonstrating that it is possible to make depressed people look either optimistic or pessimistic in the laboratory, it remains unclear what tendency they show in typical real-world judgments. As several researchers have suggested (Ackermann & DeRubeis, 1991; Dobson & Franche, 1989; Pacini et al., 1998), these studies may

not provide adequate tests of pessimistic or optimistic bias as a function of depressive symptoms. Rather, a more ecologically valid, emotionally involving task such as the prediction of future life events may provide a more valid test of the cognitive model of depression.

This study found that depressive symptoms are associated with optimistic/pessimistic bias. Evidence was consistent with the depressive bias hypothesis, but not the depressive accuracy hypothesis. As predicted by the cognitive model of depression, participants with higher levels of depressive symptoms exhibited less optimism and greater pessimism. Interestingly, our results suggest this relationship is particularly pronounced for women.

Depressive symptoms and accuracy of predictions

The association between the accuracy in participants' predictions (independent of their pessimistic or optimistic bias) and participants' severity of depressive symptoms was examined. Severity of depressive symptoms was negatively associated with accuracy in predictions (after covarying participants' pessimistic or optimistic bias). The higher the level of depressive symptoms participants reported, the less accuracy was evident in their predictions. This was true for undesirable events in particular. Thus, depressive symptoms appear to be associated with two distinct types of deficits in predictions of future life events: pessimism in predictions and imprecision in predicting undesirable events.

As already shown, depressive symptoms are associated with pessimistic or optimistic bias. Above and beyond this deficit, depressive symptoms were associated with a general inability to predict undesirable future life events with precision. To understand this effect, consider two hypothetical participants. Suppose each of these participants only gave two probability judgments for all events. One probability judgment is higher than the base rate and the second probability is (equally) lower than the base rate. However, the first of these participants gave the higher probability estimate for events that did occur and the lower probability estimate for events that did not occur. The second participant gave the higher probability estimate for events that did not occur and the lower probability estimate for events that did occur. These participants could have equal scores on Bias. However, the former participant was better able to discriminate between likely and unlikely events. In this study, it was this type of difficulty discriminating among more and less likely events (independent of pessimistic or optimistic bias) that was characteristic of participants with more severe depressive symptoms. Perhaps this difficulty reflects a general cognitive impairment in people with high levels of depressive symptoms. In the context of these findings, it would appear that it may not be appropriate to characterize the sadder as wiser (see Alloy & Abramson, 1979).

Limitations

There are several limitations of this study especially worthy of note. First, participants did not undergo diagnostic assessments as part of the study. This would have certainly provided additional, interesting information about the sample. However, this limitation is not a vital one insofar as the question of interest involves associations with depressive symptoms rather than providing a precise estimate of the extent of pessimistic bias evident among people formally diagnosed with depression.

A second limitation is that participants' reports of the occurrence of events may not have been veridical. When items were selected, the degree of subjectivity involved in evaluating whether the events occurred was considered. As much as possible, events that we thought could be rated objectively were chosen. However, it is possible that some type of systematic error may have been introduced in the assessment of whether events occurred. One's greatest concern might be that the depressive thinking of participants with greater levels of depressive symptoms would lead them to report that positive outcomes occurred less frequently than they actually did, or that low-symptom participants would report fewer negative events than what actually occurred to them. Insofar as this was the case, these more depressed participants would be prevented from exhibiting pessimistic bias (and the low-symptom participants would appear to have a less optimistic bias). Therefore, the degree of pessimistic bias associated with depressive symptoms would be underestimated. That is, insofar as depressive symptoms were associated with a tendency to underreport the frequency of positive outcomes, the estimate of the extent of pessimism observed in this study may be lower than the true level. Nonetheless, it will be important for future research to replicate this work and take additional measures to rule out possible systematic errors that may be associated with reports of whether events occurred.

The third limitation involves the events that were studied in this investigation. It was not possible to ensure that these events were randomly selected from the population of all possible desirable and undesirable events. Therefore, if several other conditions were met, this could be a limiting problem. The sample of undesirable events may have been unusual in that the frequencies of these events are generally underestimated by people, whereas people are more accurate with those events in the population that we did not sample. If this were the case, participants with higher BDI-II scores, who tended to give higher probability estimates to undesirable events, would appear relatively accurate in predicting these events. There is no reason to believe that such an unrepresentative set of events was selected, but future studies using other events could address this possibility.

A fourth limitation is that the internet-based method of data collection may have affected the results in some undetermined manner. However, available evidence for judgment and decision making research has failed to find any important differences between on-line and more traditional data collection methods that would be limiting factors in this study (Baron & Siepmann, 2000).

A final limitation concerns the lack of knowledge about characteristics of participants who reported varying degrees of depressive symptoms. For example, both state and trait pessimism may have influenced participants' BDI-II scores. While examining more specific characteristics of the sample would be informative, examining the relationship between overall severity of depressive symptoms and optimistic/pessimistic biases does provide an important test of the strong version of the depressive realism hypothesis and the cognitive model of depression. It will be important for future research to investigate whether the association between depressive symptoms and bias is accounted for by other characteristics.

Conclusion

In this study, pessimistic or optimistic bias was found to be associated with severity of depressive symptoms. This occurred in a study in which participants exhibited a wide range of

severity of depressive symptoms, an objective standard for assessing participants' judgments was available, and the judgments made by participants were real-world judgments relevant to the cognitive model of depression. Only when participants who are high-scoring on depressive symptoms are included can inferences about the nature of depression (rather than dysphoria) be made. Thus, the findings of this study suggest that the depressive accuracy hypothesis is not valid. Depressive symptoms, at a level typical of depressed people, are not associated with a lack of bias or increased accuracy in predictions. Rather, pessimistic bias and difficulty discriminating among more and less likely events (decreased accuracy) were associated with higher levels of depressive symptoms.

Appendix

Forty events investigated in this study

Average ratings from panel of 24 students		Forty events	Base rate from full sample of 153
Desirability	Controllability		
6.8	8.4	1. Will try out a new hobby, craft, or sport *	0.36
7.0	7.1	2. Will acquire a pet *	0.10
1.7	7.6	3. Will be arrested	0.01
4.1	6.9	4. Will burn something that you are cooking	0.36
7.4	2.0	5. Will find or receive a gift of a dollar or more *	0.59
2.5	3.7	6. Will have a serious headache	0.48
3.1	3.0	7. Will receive a call from a telemarketer	0.75
6.1	6.1	8. Will win in a competitive game or sport *	0.21
7.1	5.5	9. Will have an out of town friend visit you *	0.35
1.7	6.1	10. Will end a major relationship	0.08
7.1	6.5	11. Will successfully teach someone a new skill or concept *	0.53
3.1	3.5	12. Will be yelled at by stranger	0.13
7.2	7.1	13. Will learn a new skill related to work or school *	0.50
6.5	8.0	14. Will try out a new food or dish *	0.63
7.7	5.5	15. Will have a supervisor or teacher praise your work *	0.49
3.0	6.4	16. Will miss or be more than 15 minutes late for an appointment or meeting	0.34
1.9	3.4	17. Will be the victim of a crime	0.04
6.5	8.1	18. Will make a purchase in excess of $50 for your personal enjoyment *	0.56
6.5	5.0	19. Will be invited to a party *	0.54
1.6	8.5	20. Will hit someone in anger	0.03
7.2	7.1	21. Will go out of town for leisure *	0.39

2.6	8.0	22. Will bounce a check	0.05
6.8	2.5	23. Will run into an old friend that you haven't seen in a long time *	0.33
6.8	7.5	24. Will read and complete a book *	0.50
7.8	7.7	25. Will exercise at least twice a week *	0.54
6.8	8.4	26. Will write a letter or e-mail to a good friend *	0.82
1.9	8.6	27. Will be seriously ill one day because of overdrinking	0.07
1.9	6.5	28. Will have a serious family argument (not including spouse)	0.13
6.5	8.2	29. Will donate money or property to a needy person or cause *	0.38
3.1	7.8	30. Will stay up past 4 AM for school or work	0.22
2.6	6.0	31. Will have a serious disagreement with a good friend	0.13
1.8	4.8	32. Will have an injury that requires medical attention	0.05
2.2	3.9	33. Will get sick or suffer a physical illness	0.37
1.9	7.7	34. Will forget a major deadline	0.08
6.7	4.9	35. Will be asked to do an activity with a friend *	0.67
2.2	7.6	36. Will have a sexual encounter that you regret	0.05
3.1	8.1	37. Will get a parking or speeding ticket	0.04
6.9	1.6	38. Will have a relative win a contest or award *	0.10
1.4	1.2	39. Will find out that someone you know personally has died	0.18
6.7	8.1	40. Will invite a non-family member to a meal *	0.45

Note: * indicates desirable items. Other items are undesirable.

Desirability and controllability ratings were made on a 1–9 Likert-type scale with higher values indicating greater desirability and controllability.

References

Ackermann, R., & DeRubeis, R. J. (1991). Is depressive realism real? *Clinical Psychology Review*, *11*, 565–584.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*, 441–485.

Alloy, L. B., & Abramson, L. Y. (1988). Depressive realism: Four theoretical perspectives. In L. B. Alloy (Ed.), *Cognitive Process in Depression* (pp. 223–265). New York: Guilford.

Baron, J. (2000). *Thinking and deciding* (3rd ed.). New York: Cambridge University Press.

Baron, J., & Siepmann, M. (2000). Using web questionnaires for judgment and decision making research. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet*. New York: Academic Press.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Hoeber Republished as *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive therapy and emotional disorders*. New York: International Universities Press.

Beck, A. T., Riskind, J. H., Brown, G., & Steer, R. A. (1988). Levels of hopelessness in DSM-III disorders: A partial test of content specificity in depression. *Cognitive Therapy & Research*, *12*, 459–469.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the beck depression inventory-I*. San Antonio, TX: Psychological Corporation.

Carson, R. C. (2001). Depressive realism: Continuous monitoring of contingency judgments among depressed outpatients and non-depressed controls (Doctoral dissertation, Vanderbilt University, 2001). *Dissertation Abstracts International*, *62*, 1070.

Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology*, *85*, 186–193.

Coyne, J. C. (1994). Self-reported distress: Analog or ersatz depression? *Psychological Bulletin*, *116*, 29–45.

Coyne, J. C., & Gotlib, I. H. (1983). The role of cognition in depression: A critical appraisal. *Psychological Bulletin*, *94*, 472–505.

Dobson, K. S., & Franche, R. L. (1989). A conceptual and empirical review of the depressive realism hypothesis. *Canadian Journal of Behavioral Science*, *21*, 419–433.

Dobson, K. S., & Pusch, D. (1995). A test of the depressive realism hypothesis in clinically depressed subjects. *Cognitive Therapy and Research*, *19*, 170–194.

Dozois, D. J. A., & Dobson, K. S. (2001). Information processing and cognitive organization in unipolar depression: Specificity and comorbidity issues. *Journal of Abnormal Psychology*, *110*, 236–246.

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the beck depression inventory-II. *Psychological Assessment*, *10*, 83–89.

Dunning, D., & Story, A. L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology*, *61*, 521–532.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sorsky, S. M., Collins, J. F., et al. (1989). National institutes of mental health treatment of depression collaboration research program: General effectiveness of treatments. *Archives of General Psychiatry*, *46*(11), 971–983.

Farmer, A. E., & McGuffin, P. (2003). Humiliation, loss and other types of life events and difficulties: A comparison of depressed subjects, healthy controls and their siblings. *Psychological Medicine*, *33*, 1169–1175.

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *Journal of Abnormal Psychology*, *113*, 121–135.

Haaga, D. F., & Beck, A. T. (1995). Perspectives on depressive realism: Implications for cognitive theory of depression. *Behavior Research and Therapy*, *33*, 41–48.

Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology*, *100*, 555–561.

Hill, C. V., Oei, T. P., & Hill, M. A. (1989). An empirical investigation of the specificity and sensitivity of the automatic thoughts questionnaire and dysfunctional attitudes scale. *Journal of Psychopathology & Behavioral Assessment*, *11*, 291–311.

Hollon, S. D., & Garber, J. (1980). A cognitive-expectancy theory of therapy for helplessness and depression. In J. Garber, & M. E. P. Seligman (Eds.), *Human helplessness: Theory and applications* (pp. 173–195). New York: Academic Press.

Hollon, S. D., Kendall, P. C., & Lumry, A. (1986). Specificity of depressotypic cognitions in clinical depression. *Journal of Abnormal Psychology*, *95*, 52–59.

Kendall, P. C., Hollon, S. D., Beck, A. T., Hammen, C. L., & Ingram, R. E. (1987). Issues and recommendations regarding use of the Beck Depression Inventory. *Cognitive Therapy & Research*, *113*, 289–299.

Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology*, *89*, 203–212.

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin*, *101*, 259–282.

Pacini, R., Muir, F., & Epstein, S. (1998). Depressive realism from the perspective of cognitive-experiential self-theory. *Journal of Personality and Social Psychology*, *74*, 1056–1068.

Peterson, C., & Seligman, M. E. P. (1984). Causal explanations as a risk factor for depression: Theory and evidence. *Psychological Review*, *91*, 347–374.

Taylor, S. E., & Brown, J. D. (1989). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.