Depressive realism: A meta-analytic review

2 authors:

Michael T Moore
Adelphi University
46 PUBLICATIONS   **1,352** CITATIONS

SEE PROFILE

David M. Fresco
University of Michigan
133 PUBLICATIONS   **9,319** CITATIONS

SEE PROFILE

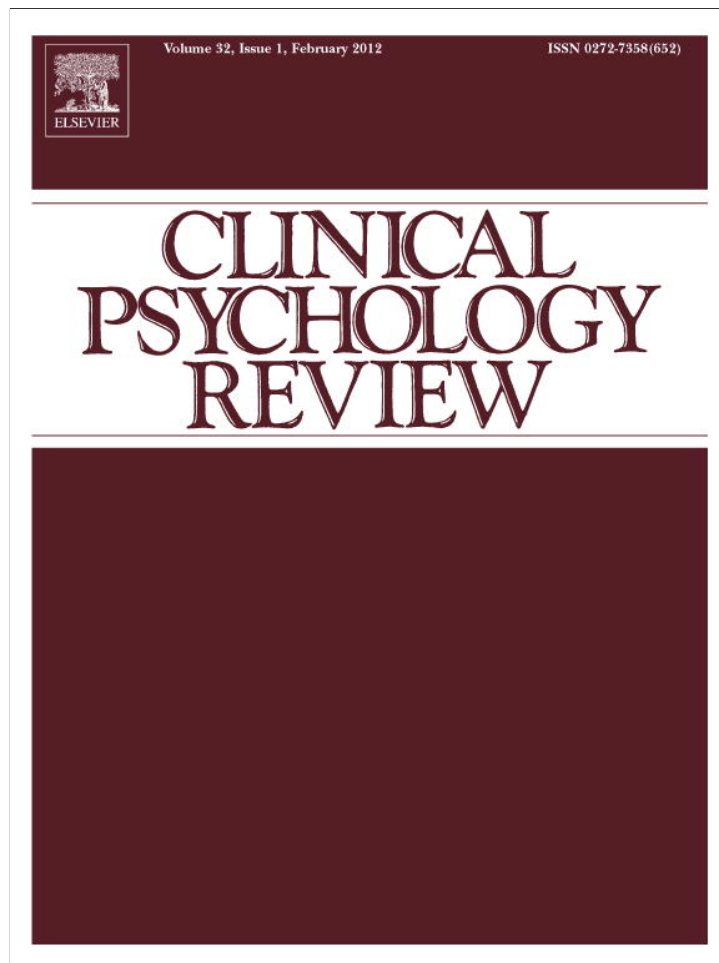Some of the authors of this publication are also working on these related projects:

Decentering as a predictor of efficacy in Mindfulness-Based Cognitive Therapy (MBCT) View project

A Comparison Between Brazilians and Americans on the Relationships Among Emotion-Regulation, Coping and Life Satisfaction. View project

CLINICAL
PSYCHOLOGY
REVIEW

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Depressive realism: A meta-analytic review

Michael T. Moore [*,1], David M. Fresco

Kent State University, Kent, OH, USA

ARTICLE INFO

ABSTRACT

The current investigation represents the first meta-analysis of the depressive realism literature. A search of this literature revealed 75 relevant studies representing 7305 participants from across the US and Canada, as well as from England, Spain, and Israel. Results generally indicated a small overall depressive realism effect (Cohen's $d = -.07$). Overall, however, both dysphoric/depressed individuals ($d = .14$) and nondysphoric/nondepressed individuals evidenced a substantial positive bias ($d = .29$), with this bias being larger in non-dysphoric/nondepressed individuals. Examination of potential moderator variables indicated that studies lacking an objective standard of reality ($d = -.15$ versus $-.03$, for studies possessing such a standard) and that utilize self-report measures to measure symptoms of depression ($d = .16$ versus $-.04$, for studies which utilize structured interviews) were more likely to find depressive realism effects. Methodological paradigm was also found to influence whether results consistent with depressive realism were found (d's ranged from $-.09$ to $.14$).

Contents

1. Introduction

Major Depressive Disorder (MDD) is a prevalent and debilitating national health problem. In the National Comorbidity Survey Replication (Kessler et al., 2003), MDD had the highest lifetime and 12-month prevalence rates (16% and 7%, respectively) of 14 major psychiatric disorders. Depression affects over 13 million individuals per year in the United States (Kessler et al., 2003). One estimate places the monetary cost in excess of $43 billion a year in treatment and lost productivity, a toll slightly larger than the costs of heart disease (Greenberg, Stiglin, Finkelstein, & Berndt, 1993). Cognitive therapy of depression (Beck, Rush, Shaw, & Emery, 1979) is one of the most empirically-validated treatments for depression (e.g., Blackburn & Moorhead, 2001; DeRubeis & Crits-Cristoph, 1998). The theory underlying cognitive therapy posits that the depressed individual is negatively biased in their perceptions, while the

* Corresponding author at: Center for Anxiety and Related Disorders, Boston University, 648 Beacon St., 6th Fl., Boston, MA 02215, USA. Tel.: +1 617 353 9610; fax: +1 617 353 9609.

   E-mail address: mtmoore@bu.edu (M.T. Moore).

[1] Now at the Center for Anxiety and Related Disorders, Boston University, Boston, MA, USA.

primary goal in cognitive therapy is returning these individuals to a more objective state (Beck et al., 1979). However, there is research that has shown that the depressed individual may be *better* able to make certain judgments than nondepressed individuals, a phenomenon referred to as "depressive realism" (see Alloy & Abramson, 1988, for a review). The literature, regarding how best to characterize the cognitions of depressed individuals, is mixed in its support. This debate calls into question how it is that cognitive therapy exerts its therapeutic effect. If depressed individuals may be *less* biased in their ability to process information than their nondepressed peers (the position of depressive realism), then how does cognitive therapy work? A recent review (Longmore & Worrell, 2007) of the literature which investigated mediators of cognitive–behavioral therapy critiqued the lack of research demonstrating that cognitive change precedes symptom change. In addition, the review highlighted research that demonstrated that symptom change in cognitive–behavioral therapy may either precede cognitive change or occur in its absence. While research has consistently demonstrated that cognitive therapy is an effective treatment for depression, knowledge of how it results in therapeutic change can result in refinements of the treatment. These refinements can potentially make cognitive therapy more concentrated, cost-effective, and hence, available to more of the millions of people who suffer from this debilitating condition. While the current study represents the first quantitative synthesis of the depressive realism literature, it is important to understand more specifically how this literature differs from the prevailing theory on the cognition of depressed individuals.

1.1. Beck's theory

Beck's (1967, 1987) theory, which formed the basis for cognitive therapy, posits that depressed affect is heavily influenced by recurrent thoughts with negative content, or automatic thoughts. These thoughts arise from deeply-held dysfunctional beliefs, or schemas, relating to the self, world, and future (e.g., "If I fail, no one will love me"). Beck identified that schemas and automatic thoughts, and the depressed affect that results from them, tend to be self-perpetuating as the depressed person both attends more to negative events in their lives and interprets events that occur after the onset of the depressed mood in light of their own dysfunctional cognitions. Beck (1987) characterizes the cognition of depressed individuals as "schema-driven" and depressed individuals themselves as possessing "depressive cognitive distortions." The thoughts of nondepressed individuals, however, are characterized as "data-driven" and he described nondepressed individuals as possessing "nondepressive accuracy," implying that depressed individuals' cognitions are systematically less informed by reality and, hence, more irrational. For instance, a depressed person may experience a significant success, but may minimize the importance of that event as due to chance because they believe that they are a failure. One of the primary goals of cognitive therapy for depression (Beck et al., 1979) is teaching depressed individuals to analytically monitor their own negative thoughts. This monitoring is done in service of both challenging and replacing these "schema-driven" thoughts with more accurate cognitions.

1.2. Depressive realism hypothesis

The "depressive realism hypothesis" (Alloy & Abramson, 1979) presented an alternative view to both conventional clinical wisdom and Beck's theory (1967, 1987) of the cognition of the depressed person. Research supportive of depressive realism illustrated not only that depressed individuals can make realistic inferences, but that they could do so to a greater extent than nondepressed individuals under certain circumstances. The first evidence for this phenomenon came in the form of studies utilizing what is called the "judgment of

contingency task." In this task, participants are asked to press a button, which results in the illumination of a light a percentage of the time that is predetermined by the experimenter. The dependent variable is the participant-rated contingency between pressing the button and the illumination of the light. As such, there are two factors that the participant needs to attend to: the occurrence of the outcome (i.e. light illumination) in the presence of the response (i.e. button press) and the occurrence of the outcome in the absence of the response. Higher positive contingencies result when the outcome occurs at a higher rate in the presence of the response than in its absence (i.e. button non-press). Negative contingencies are also possible where the outcome is less likely to occur in the presence of the response than in its absence (i.e. if pressing the button suppressed the illumination of the light). Consistent with the depressive realism effect, depressed individuals have been shown to more accurately make these kinds of judgments than nondepressed individuals (Alloy, Abramson, & Kossman, 1985; Alloy, Abramson, & Viscusi, 1981; Musson & Alloy, 1987; Vazquez, 1987). Nondepressed individuals experienced what has been referred to as an "illusion of control," where they overestimated their degree of control over the outcome. Depressed individuals experienced no such bias. In addition, these results were replicated over a variety of differing predetermined contingency conditions (Abramson, Alloy, & Rosoff, 1981; Alloy & Abramson, 1979; Dobson & Pusch, 1995; Ford & Neale, 1985; Martin, Abramson, & Alloy, 1984; Msetfi, Murphy, Simpson, & Kornbrot, 2005; Presson & Benassi, 2003; Vazquez, 1987).

Despite the number of studies utilizing the judgment of contingency task, not all of the research in support of depressive realism has used this methodological paradigm. Other methodological paradigms, referred to as self-evaluation of task performance (Gotlib, 1983; Lobitz & Post, 1979; Rozensky, Rehm, Pry, & Roth, 1977) and recall of feedback studies (DeMonbreun & Craighead, 1977; Dennard & Hokanson, 1986; Nelson & Craighead, 1977) have also produced findings compatible with depressive realism. Studies examining the self-evaluation of task performance have participants engage in a task, then rate their performance on that task without the benefit of feedback. The participants' self-performance is then compared to their actual performance to determine how accurately it was perceived. In research examining the recall of feedback, ratings of the participants' performance is given immediately after each subtask is completed, and the participants are then asked to rate their aggregate level of performance across the task as a whole. The participants' recall of the feedback they received is compared to the actual feedback to determine how accurate their recall was. In many studies (DeMonbreun & Craighead, 1977; Dennard & Hokanson, 1986; Gotlib, 1983; Lobitz & Post, 1979; Nelson & Craighead, 1977; Rozensky et al., 1977), depressed individuals were better able to evaluate or recall their performance than nondepressed individuals.

Studies comparing expectancies of success on various tasks with depressed and nondepressed individuals have replicated these findings as well (Alloy & Abramson, 1980; Alloy & Seligman, 1979; Golin, Terrel, Weitz, & Drost, 1979; Golin, Terrell, & Johnson, 1977). In many of these studies, the predictions of future success of depressed and nondepressed individuals are compared on both chance-tasks as well as tasks designed to appear skill-determined (but are actually chance-determined), both prior to and immediately after reinforcement or punishment. Smaller changes in expectancies of success by nondepressed relative to depressed individuals have been found following reinforcement or punishment in the tasks designed to appear skill-based (Alloy & Abramson, 1980; Alloy & Seligman, 1979). Insofar as performance is expected to improve on skill-determined tasks, the findings that expectancies of the nondepressed participants do not change as much as the depressed participants is taken as evidence of perceptual bias in nondepressed participants. These differences between depressed and nondepressed participants have not been found using chance-determined tasks, where

performance would not be expected to improve (Alloy & Abramson, 1980; Alloy & Seligman, 1979). Taken together, the aforementioned results have been interpreted by proponents of depressive realism as evidence that the depressed individual more accurately perceives their performance on these tasks.

1.3. Boundaries and potential functions of depressive realism

Although the above-mentioned research attests to the robustness and generalizability of the depressive realism phenomenon, there are studies that report circumstances under which depressive realism effects are not obtained (Ahrens, 1986; Alloy & Abramson, 1988; Alloy & Ahrens, 1987; Benassi & Mahler, 1985; Buchwald, 1977; DeMonbreun & Craighead, 1977; Dennard & Hokanson, 1986; Hoehn-Hyde, Schlottman, & Rush, 1982; Loewenstein & Hokanson, 1986; Moore & Fresco, 2007; Nelson & Craighead, 1977; Sacco & Hokanson, 1978, 1982; Siegel & Alloy, 1990; Tennen & Herzberger, 1987; Vazquez, 1987; Vestre & Caulfield, 1986; Wenzlaff & Berman, 1985). These boundaries, in turn, suggest how depressive realism may fit into pre-existing theory in social psychology and psychopathology. Alloy and Abramson (1988), in their comprehensive narrative review of the depressive realism literature, identified six boundary conditions on depressive realism that possessed some degree of research support. Four of these conditions refer to constraints related to situations and two refer to constraints related to the individual.

1.3.1. Situational constraints

The first of the situational constraints involves the object that is being perceived. Although the overwhelming majority of depressive realism research has asked participants to make judgments or otherwise report on their perceptions of their own behavior, some studies have compared judgments of the self versus judgments of another person between depressed and nondepressed persons (Ahrens, 1986, 1991; Ahrens, Zeiss, & Kanfer, 1988; Alloy & Abramson, 1988; Gotlib & Meltzer, 1987; Javna, 1981; Martin et al., 1984; Pyszczynski, Holt, & Greenberg, 1987; Siegel & Alloy, 1990; Vazquez, 1987). Results have shown that nondepressed participants demonstrate a positive bias in their perceptions of their own performance, but no bias in the perceptions of the performance of others. In addition, depressed participants demonstrate relatively realistic perceptions of their own performance, but a positive bias for their perceptions of others' performance (see Gotlib & Meltzer, 1987; Javna, 1981; Pyszczynski et al., 1987, and the performance of females in Martin et al., 1984 for exceptions).

The second of the situational constraints is whether the judgment or perception is made in public or private (Benassi & Mahler, 1985; Sacco & Hokanson, 1978, 1982; Strack & Coyne, 1983). Findings indicate that the cognitions of nondepressed individuals are more optimistic in public than in private, while the cognitions of depressed individuals are less responsive to the presence of others (see Strack & Coyne, 1983 for an exception to this trend).

The third situational constraint is whether the perception is made immediately or after a delay between the to-be-perceived stimulus and when the perception is assessed. Even among studies utilizing the recall of feedback paradigm, only three studies directly compared immediate perceptions to those made after a delay (DeMonbreun & Craighead, 1977; Nelson & Craighead, 1977; Wenzlaff & Berman, 1985). Both DeMonbreun and Craighead (1977) and Nelson and Craighead (1977) found that, while depressed participants' immediate perceptions were typically accurate, their memories made after a delay were negatively biased. In addition, nondepressed participants demonstrated a positive bias in both their immediate perceptions as well as their memories. Wenzlaff and Berman (1985) found similar results, with the significant exception that they found both the perceptions and memories of depressed participants to be accurate.

The final situational constraint of depressive realism is whether the to-be-perceived stimulus is ambiguous (i.e. explicitly neutral feedback) or unambiguous (i.e. clearly positive or negative feedback or information). Only one study has been conducted which has explicitly made this comparison. Dykman, Abramson, Alloy, and Hartlage (1989) evaluated the encoding of both ambiguous and unambiguous information which were both consistent and inconsistent with prior, deeply-held beliefs about the self. Results indicated that only ambiguous feedback was conducive to differential encoding by depressed and nondepressed participants.

1.3.2. Individual constraints

Alloy and Abramson (1988) also identified two constraints which involve individual factors which have some degree of research support. The first of these constraints is the severity of the depressive disorder under study. Several theorists have suggested that perceptual bias and depression may not be related in a monotonically increasing function, where degree of bias is correlated with degree of depression (e.g., Beck, 1986; Evans & Hollon, 1988; Ruehlman, West, & Pasahow, 1985). These authors have posited that nondepressed individuals may be characterized by positive biases, mildly depressed individuals by more realistic perceptions, and severely depressed individuals may be characterized by the negative perceptual and memory biases hypothesized by Beck (1967, 1976). Two studies (Dennard & Hokanson, 1986; Loewenstein & Hokanson, 1986) which have directly addressed this question have compared mildly- and moderately-dysphoric college students and both have found these groups to be equally accurate. However, McKendree-Smith and Scogin (2000) compared the perceptions of bogus, neutral personality test feedback in nondepressed, mildly, and moderately/severely depressed college students. They found that the nondepressed and mildly depressed students rated their profiles more positively than the moderately/severely depressed students. Unfortunately, this study did not address the issue of realism, per se, as it was impossible to determine which interpretation was the "correct" one, given the lack of an objective comparison (i.e., the students' actual personality profiles).

Lastly, it is possible that perceptual bias is not caused by depressed mood at all, but by some, as yet unidentified third variable(s) that is correlated with depressed mood such as self-esteem (Tennen & Herzberger, 1987, but see Crocker, Alloy, & Tabachnik-Kayne, 1988 for a failure to replicate), dysfunctional attitudes (Bynum & Scogin, 1996), or attributional style. Moore and Fresco (2007) examined the depressive realism effect in the context of a well-validated, cognitive diathesis–stress theory of the etiology of a subtype of depression, hopelessness theory (Abramson, Metalsky, & Alloy, 1989). Of interest is the finding that attributional accuracy was more closely related to attributional style (both attributional accuracy and style were measured with different instruments) than it was to symptoms of depression.

Despite the apparent wealth of findings in support of depressive realism, numerous studies have provided less favorable results. Even within the seminal Alloy and Abramson's (1979) paper in which depressive realism was first introduced, results were mixed. Some conditions (see Experiment 1) failed to produce depressive realism results altogether, while other conditions (see the noncontingency, low-density reinforcement condition in Experiment 2 and Experiment 4) failed to produce the illusion of control in nondysphoric participants. Studies assessing the accuracy of depressed and nondepressed persons' delayed recall of both task-performance (Craighead, Hickey, & DeMonbreun, 1979; DeMonbreun & Craighead, 1977) and ambiguous personality feedback (Dykman et al., 1989; Gotlib, 1983; Vestre & Caulfield, 1986) have returned results largely showing both groups to be equally accurate. The literature examining the accuracy of recall of task-performance feedback has returned consistently similar results for ambiguous feedback (Craighead et al., 1979; DeMonbreun & Craighead, 1977). Depressed individuals have been shown to underestimate positive feedback that they receive and nondepressed individuals have been shown to overestimate it (Buchwald, 1977; Wener & Rehm, 1975), illustrating bias among both groups.

1.4. Critique of the depressive realism literature

In addition to this empirical inconsistency, the methodology of some of the literature in support of a depressive realism effect has been cogently undermined. There are three primary critiques of the depressive realism literature.

Critique 1: lack of gold standard
Much of the research on the depressive realism effect has been criticized for not including a "gold-standard" of reality with which to compare participants' perceptions of events. This criticism seems to call into question the "realism" of the depressive realism hypothesis. Critiques of the depressive realism literature comes from several theorists (e.g., Ackermann & DeRubeis, 1991; Alloy & Abramson, 1988; Haaga & Beck, 1995), who perceptively note that much of the aforementioned research cannot be said to support depressive realism unequivocally as no objective standard of reality exists with which to compare many of the participants' ratings. Without a "gold-standard" measure of reality, it is theoretically impossible to state that one group or another's ratings are more or less "realistic." It should be noted that Critique 1, the lack of a gold standard of reality, regards whether or not bias can be validly assessed, not whether or not it is present. Experimental stimuli lacking a gold standard are not biased, they simply cannot be said to evaluate claims relevant to depressive realism. Bias would be demonstrated by the perceptions of a participant to stimuli that possess an objective standard of reality. In the current investigation, a study was said to possess an objective standard of reality to the extent that the stimuli, being described by the participant, could be described in an unbiased fashion at the time it was perceived. For instance, in much of the research into the expectancies of success of depressed and nondepressed persons, there is no objective standard of reality with which to compare a prediction of the future or expectancy at the time that the rating is made. Whether or not the prediction comes to pass is the "objective standard of reality," however this cannot be known by the participant at the time the predictions are made (*before* the prediction does or does not come to pass). As a result, other interpretations of the results of the expectancy studies can be plausibly offered. Ackermann and DeRubeis (1991) give the example of a nondepressed individual who may not decrease their expectancies of success following punishment for poor performance, thereby overestimating his/her chance of success, with the expectation that practice will improve their future performance. Without knowledge of how these individuals have benefited from feedback about their performance and practice in the past, it is impossible to tell if changes in their expectancies are reasonable, or "realistic," or not. It should be noted, however, that not all research into expectancies of success fails to address this critique. Some studies asked participants to predict their success on an explicitly-labeled, chance-determined task with an objective probability of success which was readily-discernable (e.g., Alloy et al., 1981, 1985; Golin et al., 1977, 1979; Lewinsohn, Mischel, Chaplain, & Barton, 1980). The expectation that practice will improve performance on a task determined purely by chance would not apply in this case. An example of such a task would be predicting the probability of rolling a single number on the roll of a die. Studies utilizing the judgment of contingency task are also excellent examples of research that provides such a gold standard. Participants are asked to rate the contingency between pressing a button and the illumination of a light, while this contingency is objectively manipulated by the experimenter and known precisely in advance.

Studies of expectancies of success or future performance were not classified as possessing a gold standard of reality in the current investigation; however, this is not to say that these studies have not made important contributions to the study of depression. The study of expectations of future positive events has important implications for hopelessness, suicide, and risk for future episodes of depression. The issue of excluding expectancy studies given their importance to the field of depression raises the related issue of how the topics "depressive realism" and "cognitive therapy of depression" are related. It is important to recognize that these two topics are related, and not identical; part of the interest of depressive realism lies in the fact that its predictions run opposite to those of cognitive therapy of depression. However, Beck's theory is much more expansive than depressive realism. It canvasses not only the presence of cognitive and perceptual biases in the depression, but also how such biases are causal to depressive disorder, and how alleviating such biases results in alleviation of the disorder. A meta-analysis attempting to cover every study of relevance to such a theory, even if only constrained to studies using depressed samples, would be lengthy indeed. Inclusion of expectancy studies may be argued on the pragmatic grounds of their importance to the field of cognitive therapy. However, this argument conflates depressive realism and cognitive therapy of depression.[2]

Critique 2: inadequate assessment of depression
The ability of self-report measures to validly assess clinical depression has also been called into question (Kendall, Hollon, Beck, Hammen, & Ingram, 1987). Other critiques of the depressive realism literature (Dobson & Franche, 1989; Haaga & Beck, 1995) highlight the fact that most of the studies that compose this literature use self-report measures, as opposed to structured clinical interviews, to assess whether participants are "depressed" or "nondepressed." As a result, this criticism would seem to call into question whether the depressive realism phenomenon really concerns "depression" at all. Some have suggested that these individuals should be labeled as "dysphoric" or "nondysphoric" to distinguish them from the clinically depressed as clinical depression is predicated on several criteria not captured by self-report measures of depression (e.g., functional impairment; Kendall et al., 1987). In addition, self-report measures of depression are ineffective at the differential diagnosis of major depressive disorder and dysthymia, the conditions of interest, from related disorders, such as bipolar disorder. Individuals with bipolar disorder would also be predicted to score highly on self-report measures of depression while in the depressive phase of their illness. As a result, it is possible that many of the participants labeled in past studies of depressive realism may not have suffered from depression, per se. Despite the aforementioned critique, however, research which has investigated depressive realism claims in both dysphoric and clinically depressed participants (Dunn, Dalgleish, Lawrence, & Ogilvie, 2007) have found similar positive biases in both groups.

Critique 3: limited external validity
Some theorists have critiqued the use of the judgment of contingency task or other laboratory tasks to assess the realism in people's perceptions of events (Dobson & Franche, 1989; Haaga & Beck, 1995). Systematic variation in experimental findings has been noted seemingly to indicate that more robust depressive realism effects are found in less externally valid, laboratory tasks. In addition, evidence of perceptual bias in depressed participants has been

---

[2] We are thankful to an anonymous reviewer for making this comment.

found in tasks that more closely mimic the judgments people make outside of the laboratory (Dobson & Franche, 1989; Moore & Fresco, 2007). This finding implies that the depressive realism effect may merely be an artifact of a particular type of task, or constrained to laboratory tasks that do not resemble real life, and is more a methodological artifact than a clinically-useful phenomenon.

1.5. The present study

Although previous reviews of the depressive realism literature (Dobson & Franche, 1989) have attempted to resolve the empirical heterogeneity in obtained results, a largely qualitative, "vote-counting" method was used to synthesize the literature. In this method, the number of studies finding in favor of or against a particular hypothesis is tallied, and the result with the most "votes" is declared the more valid. Traditionally, vote-counting relies exclusively on statistical significance and therefore ignores the size of the effects obtained in various studies. As a result, it has been criticized as more likely to result in biased conclusions than those based on more quantitative methods of research synthesis (Bangert-Drowns, 1986; Glass, McGaw, & Smith, 1981). In addition, no previous attempt to review the depressive realism literature has accounted for the three critiques mentioned above. The current study sought to quantitatively synthesize the literature on depressive realism with the hopes of resolving the empirical heterogeneity of findings while at the same time addressing the three aforementioned critiques.

Hypotheses:

1. Consistent with expectations from the depressive realism hypothesis, effects averaged across studies will show less perceptual/attentional bias in dysphoric/depressed versus nondysphoric/nondepressed participants.
2. Examination of the direction of bias in dysphoric/depressed and nondysphoric/nondepressed groups in isolation from one another will indicate that nondysphoric/nondepressed individuals will be biased toward positive stimuli, whereas dysphoric/depressed individuals will not evidence any such bias (consistent with the findings of depressive realism).
3. Studies that utilize an objective standard of reality will evidence larger depressive realism effects than studies that do not (see Critique 1).[3] No research has yet been conducted which has quantitatively evaluated the impact of this variable on the depressive realism effect. As a result, this hypothesis is largely exploratory. However, it is felt that Critique 1 is the most theoretically substantive of those listed above and has been included for this reason.
4. Method of assessment will serve as a moderator of the depressive realism effect (see Critique 2). Specifically, studies that utilize structured clinical interview will produce larger depressive realism effects than studies that utilize self-report, as it is thought that the former will result in more homogenous depressed/nondepressed groups (thereby increasing resulting effect sizes).
5. The external validity of the study will serve as a moderator of the depressive realism effect (see Critique 3). Dobson and Franche (1989) noted that much of the support for depressive realism came in the form of studies utilizing paradigms which do not well-represent perception outside of the laboratory (e.g., the judgment of contingency task). Studies which lack external validity would be expected to make this sacrifice at the expense of increase internal validity. We would expect that this increased control for extraneous variables

would result in reduced error variance and larger depressive realism effects. As a result, it is expected that studies that lack external validity will produce larger differences between dysphoric/depressed and nondysphoric/nondepressed individuals and, therefore, larger depressive realism effects.

Although it would have been ideal to evaluate the validity of the six boundary conditions on depressive realism mentioned above, several factors prevented these analyses from being statistically and methodologically feasible. For the self- versus other-reference and public versus private conditions, the majority of the research conducted does not adequately address Critique 1. Most of the authors investigating the perception of self versus other were primarily interested in relative differences on this variable. As a result, establishing which version of the percept was "right" (self or other) was not a primary aim of this research. With regard to the literature evaluating the depressive realism effect in public versus private conditions, only three studies have been conducted making this comparison. Of these three studies, only two studies addressed Critique 1 and, of these two studies, information necessary to be useful in this meta-analysis could not be obtained for one of them. A similar lack of literature prevented the examination of ambiguous versus unambiguous stimuli and severity of depression. With regard to the examination of ambiguous versus unambiguous stimuli, only one study was found. Two studies have examined the relationship between severity of depression and the depressive realism effect. However, only one of these studies adequately addresses Critique 1 and, lamentably, information necessary to be useful in this meta-analysis could not be obtained from it. While sufficient number of studies have been conducted using both immediate and delayed perceptions, this hypothesis would be almost entirely redundant with a comparison of the recall of feedback paradigm to other methodological paradigms. This paradigm is primarily differentiated from the self-evaluation of task performance paradigm by the delayed nature of the perception in question. Because the effects of recall could not be differentiated from the particular effects of the paradigm under which it was evaluated, a comparison of immediate and delayed perceptions was not included in the present investigation.

2. Method

2.1. Search procedure

The current investigation attempted to obtain data from as many studies relevant to depressive realism as possible. However, it was outside the scope of this study to attempt to canvass certain closely-related research areas. Studies utilizing the emotional Stroop and dot probe tasks in depressed and nondepressed individuals were not included in the current investigation. This exclusion was made on practical grounds as these studies could, and have (cf. MacLeod, Mathews, & Tata, 1986), composed their own, quite voluminous meta-analysis. The current investigation also did not examine the totality of studies examining memory biases in depression. This was done because much of the research examining memory biases does not attempt to directly evaluate depressive realism. Much of this literature attempts to demonstrate that depressed individuals preferentially recall negatively-valenced material and nondepressed individuals preferentially recall positively-valenced information. This paradigm attempts to evaluate *differences* between groups, but not the systematic *biases* that are the hallmark of depressive realism. In other words, this paradigm assumes that neither group is more biased, simply that both are biased equally under differing circumstances. In circumstances where this type of paradigm is not utilized (e.g., the recall of feedback paradigm), these studies were included.

Relevant studies were located by first conducting a search of PsycINFO using the search terms "depressive realism," "illusion of control," "cognitive distortion," and "judgment of contingency." Relevant articles were also selected via a thorough search of studies cited in already-

---

[3] It should be noted that the aforementioned hypothesis merely seeks to evaluate whether an objective standard of reality is a moderator of the depressive realism effect. It does not seek to quantitatively evaluate Critique 1. Critiques 2 and 3 argue that poor assessment and lack of external validity moderate depressive realism effects. Critique 1, on the other hand, argues that an objective standard of reality is an absolute, theoretical necessity when evaluating theory relevant to depressive realism, regardless of whether or not this variable exerts any influence on the results of said evaluation.

located articles. This latter selection method allowed for detection of unpublished sources (e.g. theses, dissertations, conference presentations) that are more likely to report results which are not statistically significant; addressing the so-called "file drawer problem" (Rosenthal, 1979). Relevant articles were defined as any study that: (1) could be said to examine perceptual accuracy and (2) did so via comparison of groups of depressed/dysphoric and nondepressed/nondysphoric participants. With regard to the former criteria, we used the rather liberal threshold of any study that purported to investigate bias, as defined in a non-relativistic manner. This criterion primarily excluded studies which examined perceptual differences and made no claims about accuracy, such as the memory bias studies described above.

2.2. Coding procedure

Each study was coded as to: whether the dependent variable(s) could be compared to an objective standard of reality, how depression was assessed, which methodological paradigm was used, and the degree to which this method was externally valid. Studies where the dependent variable used to assess realism was compared to an objective standard of reality (and, therefore, addressed Critique 1) were compared to studies that did not utilize an objective standard of reality to determine what influence this potential moderator variable has on the magnitude of depressive realism results obtained. Studies which examined the differential expectancies or predictions of future success on a skill-determined task(s) by depressed and nondepressed individuals did not utilize an objective standard of reality. As a result, these studies did not address Critique 11 and were categorized accordingly. Studies, which compared self-perceptions to the perception of others in depressed and nondepressed participants (i.e., without attempting to determine if either of these perceptions were more realistic or objective) were also coded as not having addressed Critique 1. As mentioned previously, studies which do not address Critique 1, and do not possess an objective standard of reality, cannot be said to evaluate depressive realism, unequivocally. Insofar as studies which do not address Critique 1 are not directly relevant to the depressive realism literature, only studies which addressed Critique 1 were utilized in the evaluation of our hypotheses (with the obvious exception of Hypothesis 3).

To address Critique 2 (that depressive realism studies really assess dysphoria instead of depression), studies that address Critique 1 were coded as to how depression was assessed. Studies that utilized clinical interview were compared to studies that, instead, utilized just self-report, to determine whether method of assessment of depression served as a moderator variable of the depressive realism effect. To address Critique 3 (that the depressive realism effect may not be replicable outside of the laboratory), studies that satisfy Critique 1 were also coded on the degree of external validity present in the dependent variable (High versus Low) to determine the influence of this moderator. Studies where the experimental task closely mimicked judgments made outside of the laboratory would be rated "High." How a study was coded was a function of both aspects of the context and methodology (stimuli presented via computer versus interaction with a confederate) as well as the nature of the variable itself. In the case of a participant asked to judge their performance on a task in the presence of objective feedback, is the task one that the participant would be likely to encounter outside of the experiment? An example of a research design that was coded as high in external validity is a study that used the participants' ratings of their performance in a social interaction that they were not informed was part of the study. An example of a research design that was coded as low in external validity was the judgment of contingency paradigm. In addition to addressing these three critiques, the experimental methodology used in a particular study was coded (judgment of contingency, recall of feedback, and evaluation of performance) to determine the potential of this variable as a moderator of the depressive realism effect. All studies submitted

to statistical analysis in the current investigation ($n = 75$) were coded by three trained raters. Raters coded practice articles until their ratings were determined to match those of a criterion coder (the first author). Adequate inter-rater reliability was obtained for whether the study possessed an objective standard of reality (intraclass correlation [ICC] = .87), method of assessment (ICC = .88), methodological paradigm (ICC = .91), and the degree to which this method was externally valid (ICC = .87).

2.3. Statistical procedure

For studies coded as addressing Critique 1, all mean raw subjective scores for both dysphoric/depressed and nondysphoric/nondepressed groups were subtracted from the objective scores, which were provided in the text of the studies themselves. Therefore, a score of zero indicates purely objective responding, while increasingly negative scores indicate negative bias, and increasingly positive scores indicate positive bias. For example, a group whose mean judgment of contingency score on the judgment of contingency task was 40, when the experimenter-determined contingency was 75, would have a mean difference score of −35. The scores' negative sign indicates that the judgment of this event was more negative than the event itself, while the absolute value indicates the degree of bias. For studies that assessed the accuracy of participants' recall for positive and negative stimuli (making a score of 100% indicate perfect accuracy), the mean scores for negative stimuli were subtracted from the means for positive stimuli. This difference score was used to make the results of all studies interpretable in the same manner, as a score of zero would indicate evenhanded accuracy, increasingly negative scores would indicate preference for negative stimuli (demonstrating a negative bias), and increasingly positive scores would indicate preference for positive stimuli (demonstrating a positive bias). For example, a group that recalled negative stimuli correctly an average of 50% of the time and positive stimuli an average of 20% of the time would have a difference score of −30, indicating a preference for recalling negative stimuli. Effect size statistics (Cohen's d) were then computed by subtracting the absolute value of the nondysphoric/nondepressed groups' scores from those of the dysphoric/depressed group, and then dividing by the pooled standard deviation.

The d statistic has been critiqued for being a biased estimator of the true population effect size in smaller samples (Hedges, 1981), therefore, a correction factor was applied according to the suggestions of Hedges (1981). Note that all descriptive statistics listed in the current investigation have been corrected for this sampling error. Using this corrected d statistic, a small effect size would indicate that both groups were equally accurate in their perceptions, while increasing positive effects indicate relatively higher degrees of accuracy in the nondysphoric/nondepressed group (contrary to predictions of depressive realism) and increasingly negative effect sizes indicate relatively higher degrees of accuracy in the dysphoric/depressed group (consistent with predictions from depressive realism).

If a single study possessed multiple, relevant dependent variables, a weighted average was computed (composed of the effect sizes of the dependent variables within a study). This was done to address critique that studies that use multiple effect size statistics (data points) from a single study violate the independence of observation assumptions of much of inferential statistics (Glass et al., 1981; Rosenthal, 1991). Weighted averages were computed using the random-effects procedure outlined by Hedges and Vevea (1998). Random-effects analyses, unlike fixed-effects analyses, do not assume that population parameters are invariant across studies (e.g., Hedges & Vevea, 1998; Schmidt, Oh, & Hayes, 2009). As a result, findings from random-effects analyses can be more readily generalized to participants that were not included in the studies being analyzed. However, the trade-off of this increased external validity is the decreased power of these statistics. Circumstances where population parameters would be presumed to vary across

studies include (but are not limited to) where an unmeasured moderator variable is present in the collection of studies or if measurement error is highly variable across studies (Schmidt et al., 2009). The use of fixed-effects analyses has been criticized on the grounds that most meta-analyses fall into at least one of these two circumstances and are concerned with generalization to studies not included in the meta-analysis itself (e.g., Field, 2003; Hedges, 1994; Hedges & Vevea, 1998; Hunter & Schmidt, 2000; Raudenbush, 1994; Rosenthal, 1991). The goal of the current investigation is to detect the presence of moderators, not all of the studies relevant to depressive realism could be included (making generalization to these studies a significant strength), and significant variability in measurement precision was observed across studies (see below). Given these three conditions, and evidence that suggests that erroneously narrow confidence intervals and inflated Type I Error results from the inappropriate use of fixed effects analyses (e.g., Field, 2003; Hedges, 1994; Hedges & Vevea, 1998; Hunter & Schmidt, 2000; Raudenbush, 1994; Rosenthal, 1991; Schmidt et al., 2009), random effects analyses were used exclusively in the current study.

Calculating effect sizes by subtracting nondysphoric/nondepressed group scores from those of the dysphoric/depressed group provided an index of the degree of perceptual accuracy in one group relative to the other. However, this approach does not provide information on how each group is biased in an absolute sense, positively or negatively. One-sample *t*-tests were computed from the signed difference scores mentioned above for the dysphoric/depressed and nondysphoric/ nondepressed groups individually, which were then converted into corrected effect size statistics (Cohen's *d*) using a supplemental formula.[4] A score of zero would indicate purely objective responding, while increasingly negative scores indicate negative bias, and increasingly positive scores indicate positive bias.

To evaluate the presence of a moderator random-effects analyses were again used. The *Q*-statistic (Hedges & Olkin, 1985) was utilized as the random-effects statistic and is used as an indication of the degree of variability in effect sizes in meta-analysis. Similar to the *F*-statistic in ANOVA, *Q*-statistics are calculated to provide estimates of the degree of variance within the levels of the moderator (Q_W) as well as between them (Q_B). To evaluate the presence of publication bias, we utilized Rosenthal's (1979) "File Drawer" Test or Fail Safe *N* (FSN). The results of this test provide an indication of the number of studies demonstrating statistically nonsignificant results that would have to exist in "file drawers" to reduce a particular effect to nonsignificance. Whenever a mean effect size is presented below, we have also included the value for the FSN.

2.4. Studies

A total of 121 studies were located that were relevant to the depressive realism literature and made at least one comparison between dysphoric/depressed and nondysphoric/nondepressed groups. Of the 121 total studies, 46 studies (38% of the total) did not provide sufficient information for effect size statistics to be calculated. These 46 studies fell into two types: (1) the authors could not be contacted to provide the missing information (13 studies, 28% of studies with insufficient information) or (2) the authors no longer possessed such information (33 studies, 72% of studies with insufficient information). The large number of missing studies is an unfortunate consequence that much of the depressive realism literature was conducted in response to the Alloy and Abramson's (1979) manuscript, which was prior to the advent of

personal computers, and the ease of data storage and retrieval that resulted from their use. For studies that did not possess sufficient information to calculate effect size statistics (including the unpublished sources mentioned above), multiple attempts were made to contact any and all authors for whom contact information could be obtained. Of the 75 studies remaining (see Table 1 for a complete list of studies and effect sizes), 36 studies (48%) addressed Critique 1. Of these 36 studies, 15 studies utilized the judgment of contingency task (42%), 12 studies utilized the recall of feedback paradigm (33%), and 9 asked the participants to make evaluations of their performance (25%). The 36 studies that addressed Critique 1 comprise 4108 participants (approximately 66% female) from across the US and Canada, as well as from England, Spain, and Israel. Unfortunately, data on age and race were provided for such a small number of studies that it precluded examination of these variables.

3. Results

In any research endeavor involving inferential statistics, random sampling is an important prerequisite in making generalizations from the particular participants sampled to the population about which the researcher wishes to draw conclusions. In meta-analysis, studies themselves, rather than participants, are the unit of analysis. Therefore, random sampling involves randomly sampling studies from the population of all relevant research articles. The difficulty in random sampling in meta-analysis lies in the tendency for studies with nonsignificant effects, dissertations, unfinished conference presentations, etc., not to be published; the so-called File Drawer Effect or publication bias. This makes sophisticated tests of publication bias a necessity in meta-analysis. To test for publication bias, we utilized the Duval and Tweedie (2000a, 2000b) trim and fill procedure. In this method, the inverse of study variances are plotted on the y-axis (such that increasing values indicate decreasing variances), while the corrected study-level effect size is plotted on the x-axis (see Figs. 1 and 2). Lack of symmetry in the plots is indicative of publication bias. For example, an asymmetrical plot due to a truncated right tail would indicate a lack of studies with larger effects. Analysis of the plot of all 75 studies (see Fig. 1) revealed that only 4 studies needed to be trimmed to correct for publication bias. Fig. 1 is left-skewed and is illustrative of a lack of studies with large variances and large effects that are contrary to depressive realism. The trim and fill procedure can also be used to correct for this publication bias where it is detected by mirror-reflecting outliers, adding this projected data, and recalculating relevant means. This is done iteratively, beginning with the largest outliers, until the corrected plot does not indicate publication bias. Correcting for these 4 studies changed the mean effect size from −.09 to −.07. The trim and fill procedure indicated that no studies are needed to be trimmed to correct for publication bias in the subset of data addressing Critique 1, the data upon which almost all of our analyses were conducted (see Fig. 2).

Hypothesis 1 predicted that dysphoric/depressed participants would illustrate a smaller degree of bias than nondysphoric/nondepressed participants. Consistent with expectations from the depressive realism hypothesis, dysphoric/depressed individuals illustrated a smaller degree of bias than nondysphoric/nondepressed individuals (weighted mean $d = -.07$, $SD = .46$, FSN = 4283). However this finding was below Cohen's (1992) convention for a small effect. In addition, the large standard deviation suggests that this mean result might not adequately characterize a substantial portion of the total literature. The results obtained using effect sizes calculated by subtracting the nondysphoric/nondepressed group mean from the dysphoric/depressed group mean, while useful, can only speak to the amount of bias that dysphoric/depressed individuals possess relative to nondysphoric/nondepressed individuals and does not address if the perceptions of either group are biased in an absolute sense. Hypothesis 2 addressed this point and posited that while dysphoric/depressed

---

[4] We are thankful to Larry Hedges for the following, helpful information (L. Hedges, personal communication, August 30, 2007). To obtain Q_B, it is necessary to calculate a constant *c*, which is defined as $1-(3/(4m-1))$, where $m =$ the degrees of freedom in the standard deviation. In the case of computing *c* with only one sample, $m = n-1$ and, therefore, $c = 1-(3/(4n-5))$. Similarly, *v* also needs to be similarly adjusted for use with one sample. The adjusted formula is, as follows: $v = (1/n)+(d^2/2n)$.

Table 1
Effect sizes, sample weights, and coded variables of depressive realism studies.

Study	Critique I addressed?	Study type[a]	Depressed/ dysphoric[a]	External validity[a]	Avg. d	w	n (% female)	d_d[b]	d_{nd}[b]
Abramson, Alloy, & Rosoff (1981)	Y	JOC	Dys	Low	.49	37.99	80 (50)		-.19
Alloy & Abramson (1979)	Y	JOC	Dys	Low	-.32	65.09	288 (50)	.93	-1.77
Blanco, Matute, & Vadillo (2009)	Y	JOC	Dys	Low	-.97	13.99	66 (NA)	2.07	3.46
Bryson, Doan, & Pasquali (1984)	Y	JOC	Dys	Low	.45	60.48	64 (50)	2.05	1.79
Carson, Hollon, & Shelton (2010)	Y	JOC	Dep	Low	.17	174.37	80 (NA)	-.83	-.23
Cobbs (1990)	Y	JOC	Dys	Low	-.09	119.33	48 (62)	.32	.24
Dobson & Pusch (1995)	Y	JOC	Dep	Low	-.02	7.50	30 (100)	3.49	2.03
Ford & Neale (1985)	Y	JOC	Dys	Low	.11	42.82	60 (NA)	.47	.09
Kapci & Cramer (1999)	Y	JOC	Dys	Low	.10	119.05	80 (51)	.31	.11
Martin, Abramson, & Alloy (1984)	Y	JOC	Dys	Low	-.20	75.98	108 (50)	3.06	3.14
Mikulineer, Gerber, & Weisenberg (1990)	Y	JOC	Dys	Low	-.12	120.54	64 (66)	.09	.83
Msetfi, Murphy, & Simpson (2007)	Y	JOC	Dys	Low	-.52	205.38	195 (53)	6.98	7.19
Msetfi, Murphy, Simpson, & Kombrot (2005)	Y	JOC	Dys	Low	-.11	108.46	224 (50)	5.38	6.04
Presson & Benassi (2003)	Y	JOC	Dys	Low	-.06	51.19	102 (100)	-.46	.35
Vazquez (1987)	Y	JOC	Dys	Low	-.21	70.86	92 (100)	.91	.72
Craighead, Hickey, & DeMonbreun (1979)	Y	ROF	Dys	Low	-.70	4.94	21 (100)	3.27	6.31
Derry & Kuiper (1981)	Y	ROF	Dys	Low	-.20	46.68	32 (100)	-1.54	.95
Dobson & Shaw (1981)	Y	ROF	Dys	Low	.39	77.58	40 (NA)	-.30	.06
Dykman, Abramson, & Albright (1991)	Y	ROF	Dys	Low	.38	44.97	92 (60)	21.95[b]	23.08[b]
Gotlib (1981)	Y	ROF	Dep	Low	.81	13.39	35 (53)	-.05	-.38
Gotlib (1983)	Y	ROF	Dep	High	.43	8.49	35 (54)	-1.57	.81
Javna (1981)	Y	ROF	Dys	High	-.33	180.28	186 (52)	.01	-.05
Johnson, Petzel, Hartney, & Morgan (1983)	Y	ROF	Dys	High	.06	18.82	40 (50)	9.27	9.64
Nelson & Craighead (1977)	Y	ROF	Dys	Low	.29	49.35	56 (NA)	.22	.16
Puseh, Dobson, Ardo, & Murphy (1998)	Y	ROF	Dys	Low	.11	39.17	79 (57)	-1.42	-1.19
Roth & Rehm (1998)	Y	ROF	Dep	Low	-.27	27.14	40 (0)	.17	.83
Wenzlaff (1984)	Y	ROF	Dys	High	-.05	647.67	358 (48)	.01	.16
Beyer (2002)	Y	EOP	Dys	High	.21	330.80	997 (62)	-.35	-.10
Bruce & Arnett (2004)	Y	EOP	Dys	High	-.29	22.23	45 (NA)	.28	.97
Bynum & Scogin (1996)	Y	EOP	Dys	High	.09	41.95	56 (NA)	-.87	.96
Dunn, Dalgleish, Lawrence, & Ogilvie (2007)	Y	EOP	Dys	Low	.88	20.53	90 (69)	.46	.78
Harkness, Sabbagh, Jacobson, Chowdrey, & Chen (2005)	Y	EOP	Dys	High	-.42	9.84	43 (100)	-2.48	-2.81
Johnson & DiLorenzo (1998)	Y	EOP	Dys	High	-.25	16.12	72 (50)	.17	.66
Moretti (1985)	Y	EOP	Dys	High	-.80	22.13	100 (51)	.61	1.86
Stone, Dodrill, & Johnson (2001)	Y	EOP	Dys	Low	-.10	38.10	83 (NA)	.06	.77
Strack & Coyne (1983)	Y	EOP	Dys	High	-.42	44.05	120 (100)	1.26	-1.08
Abramson, Garber, Edwards, & Seligman (1978)	N	EXP			-1.73	2.91	16 (63)		
Ahrens (1991)	N	SOC			-.03	81.35	114 (NA)		
Ahrens, Zeiss, & Kanfer (1998)	N	SOC			-.28	60.60	73 (38)		
Andersen (1990)	N	EXP			-2.54	6.93	82 (60)		
Cane & Gotlib (1985)	N	EXP			.21	71.29	48 (69)		
Crocker, Alloy, & Tabachnik-Kayne (1988)	N	SOC			.28	62.83	45 (59)		
DeMonbreun & Craighead (1977)	N	OTR			.32	46.47	32 (0)		
Dunning & Story (1991)	N	EXP			-.43	70.33	423 (NA)		
Dykman, Abramson, Alloy, & Hartlage (1989)	N	OTR			.28	40.23	84 (64)		
Dykman, Horowitz, Abramson, Usher (1991)	N	SPF			-.42	85.32	120 (50)		
Finkel, Glass, & Merluzzi (1982)	N	OTR			-.02	142.86	60 (0)		
Garber & Hollon (1980)	N	EXP			-.10	64.07	66 (50)		
Glass, McKnight, & Valdimarsdottir (1993)	N	OTR			-2.33	21.41	162 (93)		
Gotlib (1982)	N	SPF			-.10	15.43	162 (93)		
Gotlib & Meltzer (1987)	N	SPF			-.51	28.88	40 (100)		
Hammen & Krantz (1976)	N	OTR			-.61	90.90	67 (100)		
Hancock, Moffoot, & O'Carroll (1996)	N	OTR			-.38	12.92	28 (57)		
Kapci & Cramer (1998)	N	EXP			.09	48.92	58 (55)		
Klein (1975)	N	EXP			.27	14.88	64 (66)		
Krantz & Gallagher-Thompson (1990)	N	OTR			.09	119.67	62 (68)		
Loeb, Beek, & Diggory (1971)	N	EXP			-3.59	4.89	40 (0)		
Loewenstein & Hokanson (1986)	N	SPF			-1.32	19.71	51 (100)		
Lovejoy (1991)	N	SPF			-.55	15.34	32 (100)		
Margo, Greenberg, Fisher, & Dewan (1993)	N	OTR			-.63	77.89	314 (32)		
McKendree-Smith & Scogin (2000)	N	OTR			.01	9.24	19 (64)		
McNamara & Hackett (1986)	N	SPF			-.45	54.57	239 (70)		
Miller & Seligman (1973)	N	EXP			-.65	15.02	32 (41)		
Miller & Seligman (1976)	N	EXP			-.32	34.81	48 (52)		
Miller, Seligman & Kurlander (1975)	N	EXP			-1.15	6.65	31 (NA)		
Pacini, Muir, & Epstein (1998)	N	OTR			-.48	161.89	75 (52)		
Pyszezynski, Holt & Greenberg (1987)	N	EXP			-.61	31.81	108 (100)		
Rosenfarb, Burker, Morris, & Cush (1993)	N	EXP			.99	21.25	24 (100)		
Sacco & Hokanson (1978)	N	EXP			.26	7.93	32 (75)		
Sacco & Hokanson (1982)	N	OTR			.10	40.32	44 (64)		
Stone & Glass (1986)	N	OTR			.06	95.05	51 (73)		
Strunk & Adler (2009)	N	EXP			.77	23.37	85 (64)		
Strunk, Lopez, & DeRubeis (2006)	N	EXP			.63	25.45	122 (70)		
Vestre & Caulfield (1986)	N	OTR			-.37	16.72	35 (51)		
Whitton, Larson, & Hauser (2008)	N	EOP			.65	8.14	133 (50)		

Note: Average weighted effect across all studies = -.10; Statistical analyses were only conducted on this variable for studies which addressed Critique 1; This variable was only coded for studies which were evaluated for the presence of moderator variables (i.e, addressed Critique 1); Avg.d = Average effect size (Cohen's *d*) comparing perceptual accuracy of depressed/dysphoric and nondepressed/nondysphoric participants; w = Inverse variance weight; JOC = Judgment of Contingency; ROF = Recall of Feedback; EOP = Evaluation of Performance; EXP = Expectancies of Success; SOC = Social Comparison; OTR = Other; d_d = Average d for depressed participants; d_{nd} = Average d for nondepressed participants; [b] = Outlier not included in analysis; NA = Gender information not available; Small effect = .20, Medium effect = .50, Large effect = .80.
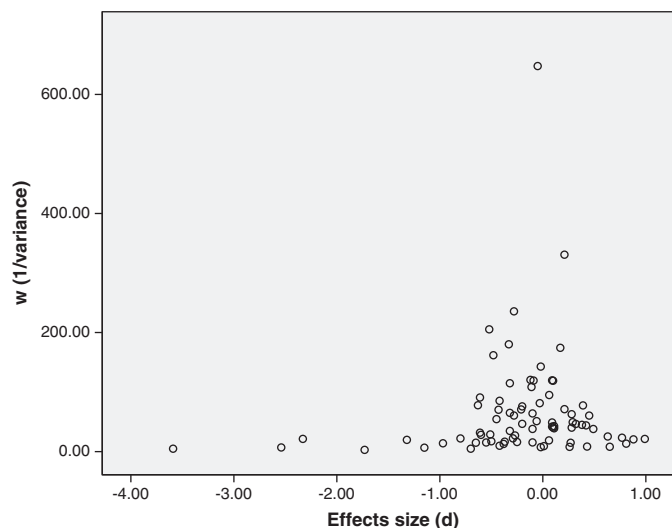
Fig. 1. Funnel plot of all effect sizes as a function of inverse-variance weights (w). N = 75.

participants would not evidence any significant bias, nondysphoric/nondepressed participants would demonstrate a bias for positively-valenced stimuli. Analyses that examined each group individually indicated that individuals in the dysphoric/depressed group tended to be biased optimistically (weighted mean $d = .14$, $SD = 2.42$, FSN = 8347), however, this result was less than a small effect. Nondysphoric/nondepressed individuals also illustrated an optimistic bias, although to a greater extent (weighted mean $d = .29$, $SD = 2.53$, FSN = 4777), and exceeded the convention for a small effect. Given these findings, Hypothesis 2 can be said to be partially supported. The findings that nondysphoric/nondepressed individuals evidence a larger degree of absolute bias and are biased positively, are consistent with both Hypothesis 2 and the expectations of depressive realism. However, that both groups demonstrate a positive bias is not consistent with it. In addition, the large variability present here both requires caution in over-interpreting the above results and suggests the presence of moderator variables, which are discussed below.

Hypothesis 3 stated that studies that utilized an objective standard of reality, and thereby adequately addressed Critique 1, would produce larger effect sizes than studies that did not. This variable did serve as a significant moderator of the depressive realism effect (Q_B [$df = 1$] = 6.87, $p = .0088$, k [number of studies] = 75, total $n = 7305$). Examination
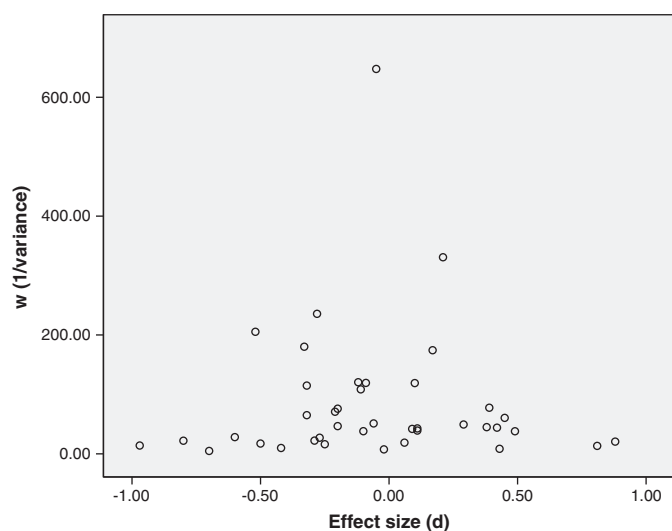

Fig. 2. Funnel plot of effect sizes for studies which address Critique 1 as a function of inverse-variance weights (w). N = 36.

of average effects for both studies adequately addressing Critique 1 (weighted mean $d = -.03$, $SD = .41$, FSN = 868) as well as those studies that did not (weighted mean $d = -.15$, $SD = .51$, FSN = 1245) indicated that both types of studies found depressive realism effects. However, counter to expectations, this effect was much stronger in studies lower in methodological quality.

Hypothesis 4 stated that studies that utilize structured clinical interview would produce larger effects than studies that utilize self-report. Method of assessment influenced whether depressive realism effects were found (Q_B [$df = 1$] = 7.57, $p = .0059$, $k = 36$, total $n = 4108$). Contrary to prediction, studies utilizing self-report were more likely to find depressive realism effects (weighted mean $d = -.04$, $SD = .40$, FSN = 717) than those that utilized structured clinical interview (weighted mean $d = .16$, $SD = .48$, FSN = 10).

Finally, Hypothesis 5 stated that studies that more readily generalized outside of the laboratory would be less likely to produce depressive realism effects. Results indicated that the external validity of the study did serve as a significant moderator of the depressive realism effect (Q_B [$df = 1$] = 32.80, $p < .0001$, $k = 36$, total $n = 4108$). Contrary to predictions, the weighted average effect size for studies low on external validity was almost identical to studies high on external validity (weighted mean $d = -.03$, $SD = .38$, FSN = 76, and weighted mean $d = -.02$, $SD = .48$, FSN = 357, respectively).

Exploratory analyses were conducted using methodological paradigm (judgment of contingency, recall of feedback, and evaluation of performance) as a moderator of the depressive realism effect as no prior research has been done on this topic and there was little theory available to guide the formation of specific hypotheses. Methodology type was found to be a significant moderator (Q_B [$df = 2$] = 19.10, $p = .00007$, $k = 36$, total $n = 4108$) and the results were, therefore, decomposed further via examination of the size of effects associated with the four major methodological paradigms used in the depressive realism literature. Both relative bias (effect sizes calculated using dysphoric/depressed and nondysphoric/nondepressed groups) and absolute bias (one-sample t-tests converted to effect size statistics) were examined. Surprisingly, results from studies using the judgment of contingency task only demonstrated a small overall depressive realism effect (weighted mean $d = -.09$, $SD = .37$, FSN = 96), despite the fact that this was the paradigm in which the depressive realism effect was first demonstrated (Alloy & Abramson, 1979). Examination of depressed and nondepressed participants separately, to determine the degree of absolute bias, indicated that both depressed and nondepressed participants overestimated the degree of contingency to the same extent (weighted mean $d = .53$, $SD = 2.26$, FSN = 421 and .60, $SD = 2.32$, FSN = 92, respectively). Both of these results exceed the convention for a large effect. The depressive realism effect was examined using this paradigm with varying degrees of objective contingency between pressing the button and the onset of the light (-50% to 100%; negative contingencies represent button pressing resulting in the suppression of illumination). Exploratory analyses were conducted evaluating degree of contingency as a potential moderator of the depressive realism effect. Unfortunately, the small number of judgment of contingency studies using each of these degrees of contingency individually necessitated aggregation into groups. Therefore, contingency was separated into Low (-50–49%) and High (50%–100%) groups. This division created groups with roughly equal numbers of studies (Low = 14 studies; High = 9 studies).[5] Interestingly, studies using a low pre-determined contingency produced results that are more consistent with depressive realism (weighted mean $d = -.20$, $SD = .71$, FSN = 84) than studies using a high contingency (weighted mean $d = .03$, $SD = .42$, FSN = 7).

---

[5] The total number of studies using low and high degrees of contingency (21) is greater than the total number of judgment of contingency studies (14) because many studies evaluated multiple degrees of contingency. However, in cases where one study had multiple effect sizes that fit into either category, these effect sizes were averaged so that studies did not contribute multiple data points to each group.

This difference, while seemingly small, was significant as degree of contingency did serve as a moderator of the depressive realism effect in studies utilizing the judgment of contingency task (Q_B [$df=1$] $=16.91$, $p=.00004$, $k=23$, total $n=1588$).

Results from the evaluation of performance studies indicated slightly more bias among depressed/dysphoric participants (weighted mean $d=.14$, $SD=.50$, FSN$=79$). Recall of feedback studies were more equivocal in their results (weighted mean $d=-.03$, $SD=.40$, FSN$=39$). However, the results from both of these methodological paradigms corresponded to less than small effects. Depressed participants in the evaluation of performance paradigm were relatively evenhanded and evidenced only a small negative bias (weighted mean $d=-.06$, $SD=1.07$, FSN$=252$), while the nondepressed subjects possessed a small, but positive bias (weighted mean $d=.14$, $SD=1.40$, FSN$=258$). Similar results were obtained in the recall of feedback studies where depressed participants were relatively evenhanded and evidenced only a small negative bias (weighted mean $d=-.10$, $SD=3.12$, FSN$=337$), whereas nondepressed participants possessed a small, but positive bias (weighted mean $d=.14$, $SD=3.31$, FSN$=156$).

4. Discussion

The current investigation serves as the first attempt to quantitatively summarize and investigate the depressive realism literature. Although the results averaging across all studies addressing Critique 1 were generally supportive of the depressive realism hypothesis, the magnitude of the effect was small. However, the large degree of variability in the size of the effects obtained by the various studies in the depressive realism literature ($Q_{Total}=493.89$ [$df=74$], $p<.001$, $SD=.72$, range: -3.59–$.99$) resulted in the small depressive realism effect obtained when studies were averaged. Dysphoric/depressed participants were found to be relatively evenhanded in their perceptions, while nondysphoric/nondepressed participants evidenced a more substantial positive bias. Substantial variability was also found among these groups, suggesting caution in interpreting these results, as well as the presence of moderator variables, discussed below.

The manner in which Hypothesis 2 was evaluated in the current investigation is worthy of comment. Bias was investigated both in dysphoric/depressed and nondysphoric/nondepressed groups relative to one another, as well as compared to an absolute standard. Much of the literature on depressive realism has not differentiated between these two methods of evaluating the theory. Past research (e.g., Dobson & Franche, 1989; Dunn et al., 2007) has noted the importance of assessing both perceptions in one group relative to another (relative bias) and comparing one group's perceptions to an absolute standard to reality (absolute bias). We echo their suggestion that future investigations of depressive realism specify the type of bias, relative, absolute, or somewhere between the two, that is being predicted. The type of bias assessed has important implications for the theory that is being investigated. We argue that two versions of depressive realism have been implicitly studied. We propose that the version of depressive realism which posits only relative bias be referred to as weak depressive realism. Weak depressive realism posits merely that depressed/dysphoric participants demonstrate less bias than nondepressed/nondysphoric participants. We propose that the version of depressive realism that makes more restrictive claims and posits both relative and absolute bias be referred to as strong depressive realism. This version of the depressive realism hypothesis posits both that depressed/dysphoric participants demonstrate a lack of significant positive or negative bias and demonstrate less bias than nondepressed/nondysphoric participants. Additionally, an intermediate version of depressive realism could, for instance, specify the direction of bias. For example, it could require that depressed/dysphoric participants demonstrate less bias than nondepressed/nondysphoric participants and also that the bias in both groups be positive. The current investigation is an example of where

making this differentiation is significant and not doing so could lead to confusion. Our results are partially supportive of strong depressive realism (which is the view we described in our hypotheses above), but fully supportive of an intermediate or weak version of depressive realism. However, this degree of support for depressive realism should be interpreted in light of the results of our moderation analyses.

An attempt was made to model the extent of the variability in the results of the depressive realism literature via investigation of theoretically-identified moderators of the depressive realism effect. These analyses indicated that depressive realism effects were more likely to be found in the absence of an objective standard of reality and when self-report (as opposed to clinical interview) was used to assess level of dysphoria. Analyses were also conducted that suggested that depressive realism effects were more equivocal in studies utilizing the judgment of contingency, evaluation of performance, and recall of feedback paradigms. This result, the lack of a strong depressive realism effect in the judgment of contingency paradigm, where depressive realism has been more frequently evaluated, was particularly surprising. Additional analyses were conducted attempting to model the heterogeneity in studies utilizing the judgment of contingency paradigm where it was discovered that depressive realism effects were slightly more likely when a lower experimenter-determined contingency was used. This finding is of particular theoretical interest and is relevant to conjectures researchers have made since the beginning of research into depressive realism (Alloy & Abramson, 1979; Msetfi et al., 2005). Some investigators have questioned whether the ability of depressed participants to accurately judge zero contingency conditions is the result of these conditions matching their preconceptions about their relationship to the world (Alloy & Abramson, 1979; Langer, 1975; Msetfi et al., 2005). That is to say, depressed individuals do not believe that their actions have any influence on events and it is coincidence that accuracy in a zero contingency condition corresponds to this view. However, our data are too limited to support the assertion that depressive realism results at lower levels of contingency in the judgment of contingency task are merely an artifact. For example, due to the small number of judgment of contingency studies overall ($n=15$), High and Low contingency groups had to be created using a median split. This artificial dichotomization may have resulted in the small differences observed between the two groups by creating groups that were not homogenous with regard to their performance on the judgment of contingency task. Additional research will need to be conducted to experimentally determine if this is the case. In particular, research is needed examining judgments of contingency using levels of contingency higher than zero. Of the 15 articles found which examined judgments of contingency, 13 of these used a zero contingency condition, and 5 did so exclusively.

Unfortunately, varying the nature of the "reality" that participants are asked to report upon has only been attempting using the judgment of contingency task. It is possible, for example, that depressive realism effects may be constrained to recall of feedback studies where negative feedback was given. However, the valence of feedback has never, to our knowledge, been systematically varied to determine its influence on whether depressive realism effects were obtained. Future research in depressive realism should focus on examining to what extent the match between participant schemata and "reality" may underlie the depressive realism effect, which would involve systematic variation of important aspects of "reality."

Although the aggregated results of the current investigation are certainly suggestive that the depressive realism effect may be constrained to a very particular set of circumstances, they should not be interpreted as suggesting that the depressive realism effect is not a valid phenomenon. There were only a limited number of studies utilizing clinical interview ($n=4$), externally valid stimuli ($n=11$), and particular research paradigms (average $n=12$, range 9–14). The small number of studies using externally valid research designs is an unfortunate consequence on the popularity of the JOCT among depressive realism researchers and the relative importance placed on internal

validity in the early investigations of a topic. The results of the current investigation suggest that future investigations of depressive realism should utilize structured clinical interview and externally valid designs. Future research should also attempt to collect data from participants displaying a range of severity in symptoms of depression. As mentioned previously, the potential for a curvilinear relationship between depression and perceptual accuracy has only been adequately evaluated once before (and so could not be examined in the current study). The questions of whether or not depressive realism is relevant to individuals suffering from a mood disorder, or outside of the laboratory at all, loom large. In addition, there was a statistically significant degree of unaccounted for variance within each of the levels of our moderator variables (i.e., Q_W). Future research will hopefully detect variables that can successfully model and account for this variability. However, for this increased accuracy in statistical prediction to occur, not only will more researches have to be done, but a consistent set of stimuli will have to be developed. Even within the judgment of contingency literature, which is centered on a particular task, the judgment of contingency task itself has taken many forms: from a physical button and light bulb to several forms of a computerized task. This lack of consistency in stimuli no doubt partially accounts for the vast heterogeneity in the results of judgment of contingency studies and other depressive realism studies. It is difficult to attempt to theoretically model unaccounted for variance when the variables making up that variance are not consistent from study to study.

Another factor that has been under-investigated in the depressive realism literature is the role of comorbid anxiety. Numerous studies have demonstrated rates of comorbidity that are alarmingly high (e.g., Brawman-Mintzer at al., 1993; Brown et al., 2001; Kessler et al., 2005). Given this degree of overlap, it is possible that the depressive realism phenomena might not be specific to depression or, at worst, be better accounted for by symptoms of anxiety. Dunn et al. (2009) recently evaluated depressive realism in the context of the tripartite model of mood and anxiety disorders (Clark & Watson, 1991; Clark, Watson, & Mineka, 1994). The tripartite model proposes that mood and anxiety disorders are best represented by symptom dimensions that are particular to each (i.e., low positive affect/anhedonia and anxious arousal, respectively) and a non-specific 'general distress' dimension. Dunn et al. (2009) found that positive self-judgment bias was uniquely and negatively related to symptoms of anhedonia and unrelated to the anxious arousal dimension. Future research should attempt to replicate this work and extend it by supplementing the assessment of depression and anxiety with structured interview. In addition, future research should examine the relationship between comorbid anxiety and depressive realism using many of the various research paradigms mentioned above.

The results of the current investigation also have clinically-relevant theoretical implications. These results can be used to explain the dichotomy between researchers finding statistically significant depressive realism effects while practitioners fail to notice such effects in their clients. It is possible that the depressive realism effect is not present under conditions normally encountered in therapy, as a result of the effect of some moderator variable(s). This hypothesis is supported by the fact that almost all of the levels of the various moderator variables examined above possessed a significant degree of variability. Potential moderator variables that have not been adequately investigated are whether participant responses refer to self versus other, responses made in public versus private settings, the level of ambiguity of the stimuli used, severity of depression, and the type of cognitive processing required (attention, encoding, or retrieval from memory). With regard to biases in information processing, it is possible that there are significant biases in the attention of nondepressed individuals relative to those suffering from depression. However, this bias may not only disappear, but reverse itself, in processes that occur later in information processing (during memory encoding or retrieval). Attentional bias research has found that nondepressed individuals

reliably evidence either a bias toward stimuli likely to result in a positive mood or away from stimuli likely to result in a negative mood, and that depressed individuals show no such bias (Gotlib, McLachlan, & Katz, 1988; McCabe & Gotlib, 1995; McCabe & Toman, 2000; McCabe, Gotlib, & Martin, 2000). However, research into autobiographical memory recall (Goddard, Dritschel, & Burton, 1996; Kuyken & Brewin, 1995; Kuyken & Dalgleish, 1995; Puffet, Jehin-Marchot, Timsit-Berthier, & Timsit, 1991), as well as memory recall research in general (Bradley & Mathews, 1983; Derry & Kuiper, 1981; Gilboa & Gotlib, 1997; MacLeod et al., 1986), finds that depressed individuals show a preference for negatively-valenced, self-referent information and that this bias reliably predicts the occurrence of depressive symptoms, introducing the possibility that bias in depressed individuals appears after attention, but either before or during memory recall. While biased memory recall may be more salient or noticeable therapeutically than attentional biases, it would not be surprising that clinicians do not make note of the depressive realism effect. Evidence that nondysphoric/nondepressed individuals evidence generally positively-biased perceptions relative to dysphoric/depressed individuals also lends support to the notion that perceptual bias may serve to protect an individual from the occurrence of depression (Alloy & Clements, 1992). It is also possible that the potential positivity bias present in the nondysphoric/nondepressed accounts for why depressive realism effects may not be noticed in therapy. To the extent that therapists are euthymic, they may tend to pathologize their dysphoric/depressed clients, whose outlook is so much more negative than their own.[2]

Research demonstrating the potentially protective function of positively biased perceptions also highlights another under-investigated area in depressive realism: the function of perceptual bias. While the study by Alloy and Clements (1992) frames the question in a dichotomous fashion, bias is either good or bad, it is possible that the value of a positive/negative perceptual bias may depend on the situation. If demonstrated, this possibility suggests that successful therapy could consist of *either* alteration of trait-like cognitive structures (as in traditional cognitive therapy) *or* a change in context.

The current investigation, and the depressive realism literature as a whole, is relevant to the question of how cognitive therapy results in reductions in depression. The theory behind cognitive therapy (Beck et al., 1979) suggests that the cognitions of depressed individuals are negatively biased and making these thoughts more realistic is the process by which the therapy exerts its effect. Depressive realism suggests that depressed individuals are *more* accurate in their perceptions and, by extension, calls into question how cognitive therapy works. While it would be tempting to frame the results of the current study as supporting one theory versus the other, it would be more accurate to see it as a beginning in defining the boundaries of the depressive realism phenomena. Future research will be needed to elaborate on the relevance of depressive realism for etiological models of depression (e.g., Moore & Fresco, 2007) and for therapy. Specifically, past research has utilized process measures which do not attempt to evaluate the accuracy of client thoughts (e.g., DeRubeis et al., 1990; Dimidjian et al., 2006; Jacobson et al., 1996). Future work could utilize methodologies relevant to both depressive realism and mediators of cognitive therapy, such as the one used in Moore and Fresco (2007), to evaluate if cognitive therapy works by making thoughts more realistic or just making them more positive.

Although the current investigation provides the first foray into quantitative review of a theoretically rich literature, some limitations of the current design warrant mention and effect the conclusions that can be drawn from it. First, slightly fewer than 40% of the total studies relevant to the depressive realism literature could not be obtained and submitted to analysis. A large amount of missing data is inevitable given the age of the depressive realism literature. Our analyses indicated that these missing data were randomly distributed and were not likely to significantly influence the results of our subsequent analyses. Nonetheless, it is possible that different results would be

obtained had this missing data been available. One important variable that was almost universally missing from published reports was data on race, which precluded analyses of the validity of depressive realism in different racial/ethnic groups.

Despite the aforementioned limitations, the current investigation provides insight into deficits in the depressive realism literature as a whole, as well as potentially fruitful areas in need of further investigation. Future depressive realism studies would benefit from more attention given to generalizing depressive realism effects outside of the laboratory and identifying depressed and nondepressed individuals using structured clinical interview. In addition, more research should be conducted using treatment-seeking samples. While sample characteristics were not formally under study in the current investigation, it should be noted that only 11 of the 36 studies in question utilized a treatment-seeking sample. This vastly limited the conclusions that could be drawn about the depressive realism phenomenon in this population. Finally, more research in depressive realism in general, and work in developing widely-accepted stimuli in particular, may help to model the large degree of variability in the results of depressive realism studies that were observed. The influence of other variables which may significantly alter the depressive realism effect, such as the degree of contingency in judgment of contingency studies and type of information processing tapped, should also be explored. If nothing else, the current investigation highlights that the depressive realism effect is far from universal. The question for future research then becomes: under what circumstances, and for which groups of people, is the depressive realism effect valid?

References[6]

*Abramson, L. Y., Alloy, L. B., & Rosoff, R. (1981). Depression and the generation of complex hypotheses in the judgment of contingency. *Behaviour Research and Therapy*, *19*, 35–45.

*Abramson, L. Y., Garber, J., Edwards, N. B., & Seligman, M. E. P. (1978). Expectancy changes in depression and schizophrenia. *Journal of Abnormal Psychology*, *87*, 102–109.

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*, 358–372.

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

Ahrens, A. H. (1986). Choice of social comparison targets by depressed and nondepressed students. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

*Ahrens, A. H. (1991). Dysphoria and social comparison: Combining information regarding others' performances. *Journal of Social and Clinical Psychology*, *10*, 190–205.

*Ahrens, A. H., Zeiss, A. M., & Kanfer, R. (1988). Depressive deficits in interpersonal standards, self-efficacy, and social comparison. *Cognitive Therapy and Research*, *12*, 53–67.

*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. (1980). The cognitive component of human helplessness and depression: A critical analysis. In J. Garber, & M. E. P. Seligman (Eds.), *Human helplessness: Theory and application* (pp. 59–70). New York: Academic Press.

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

Alloy, L. B., Abramson, L. Y., & Kossman, D. (1985). The judgment of predictability in depressed and nondepressed college students. In F. R. Brush, & J. B. Overmier (Eds.), *Affect, conditioning and cognition: Essays on the determinants of behavior* (pp. 229–246). Hillsdale, NJ: Erlbaum.

Alloy, L. B., Abramson, L. Y., & Viscusi, D. (1981). Induced mood and the illusion of control. *Journal of Personality and Social Psychology*, *41*, 1129–1140.

Alloy, L. B., & Ahrens, A. H. (1987). Depression and pessimism for the future: Biased use of statistically relevant information in predictions for self versus others. *Journal of Personality and Social Psychology*, *52*, 366–378.

Alloy, L. B., & Clements, C. M. (1992). Illusion of control: Invulnerability to negative affect and depressive symptoms after laboratory and natural stressors. *Journal of Abnormal Psychology*, *101*, 234–245.

Alloy, L. B., & Seligman, M. E. P. (1979). In G. H. Bower (Ed.), *On the cognitive component of learned helplessness and depression. The psychology of learning and motivation*, *Vol. 13*. (pp. 219–276) New York: Academic Press.

*Andersen, S. M. (1990). The inevitability of future suffering: The role of depressive predictive certainty in depression. *Social Cognition*, *8*, 203–228.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic methods. *Psychological Bulletin*, *99*, 388–399.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York: Harper & Row.

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

Beck, A. T. (1986). Cognitive therapy, behavior therapy, psychoanalysis, and pharmacotherapy: The cognitive continuum. In J. B. W. William, & R. L. Spitzer (Eds.), *Psychotherapy research: Where are we and where should we go?* (pp. 114–134). New York: Guilford Press.

Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy: An International Quarterly*, *1*, 5–37.

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

Benassi, V. A., & Mahler, H. I. M. (1985). Contingency judgments by depressed college students: Sadder but not always wiser. *Journal of Personality and Social Psychology*, *49*, 1323–1329.

*Beyer, S. (2002). The effects of gender, dysphoria, and performance feedback on the accuracy of self-evaluations. *Sex Roles*, *47*, 453–464.

Blackburn, I., & Moorhead, S. (2001). Depression. In W. L. Lyddon, & J. V. Jones (Eds.), *Empirically supported cognitive therapies: Current and future applications*. New York: Springer.

*Blanco, F., Matute, H., & Vadillo, M. A. (2009). Depressive realism: Wiser or quieter? *Psychological Record*, *59*, 551–562.

Bradley, B., & Mathews, A. (1983). Negative self-schemata in clinical depression. *British Journal of Clinical Psychology*, *22*, 173–181.

Brawman-Mintzer, O., Lydiard, R. B., Emmanuel, N., Payeur, R., Johnson, M., Roberts, J., Jarrell, M. P., & Ballenger, J. C. (1993). Psychiatric commorbidity in patients with generalized anxiety disorder. *The American Journal of Psychiatry*, *150*, 1216–1218.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, *110*, 585–599.

*Bruce, J. M., & Arnett, P. A. (2004). Self-reported everyday memory and depression in patients with multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *26*, 200–214.

*Bryson, S. E., Doan, B. D., & Pasquali, P. (1984). Sadder but wiser: A failure to demonstrate that mood influences judgment of control. *Canadian Journal of Behavioral Science*, *16*, 107–119.

Buchwald, A. M. (1977). Depressive mood and estimates of reinforcement frequency. *Journal of Abnormal Psychology*, *86*, 443–446.

*Bynum, J., & Scogin, F. (1996). The impact of dysfunctional attitudes on depressive realism. *Journal of Social and Clinical Psychology*, *15*, 305–317.

*Cane, D. B., & Gotlib, I. H. (1985). Depression and the effects of positive and negative feedback on expectations, evaluations, and performance. *Cognitive Therapy and Research*, *9*, 145–160.

*Carson, R. C., Hollon, S. D., & Shelton, R. C. (2010). Depressive realism and clinical depression. *Behaviour Research and Therapy*, *48*, 257–265.

Clark, L. A., & Watson, D. (1991). Theoretical and empirical issues in differentiating depression from anxiety. In A. Becker, & J. Kleinman (Eds.), *Psychosocial Aspects of Depression* (pp. 39–65). Hillsdale, NJ: Erlbaum.

Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology*, *103*, 103–116.

*Cobbs, D. L. (1990). Judgment of contingency and the cognitive functioning of clinical depressives. Unpublished doctoral dissertation, University of North Texas, Denton, TX.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

*Craighead, W. E., Hickey, K. S., & DeMonbreun, B. G. (1979). Distortion of perception and recall of neutral feedback in depression. *Cognitive Therapy and Research*, *3*, 291–298.

*Crocker, J., Alloy, L. B., & Tabachnik-Kayne, N. T. (1988). Attributional style, depression and perceptions of consensus for events. *Journal of Personality and Social Psychology*, *54*, 840–846.

*DeMonbreun, B. G., & Craighead, W. E. (1977). Distortion of perception and recall of positive and neutral feedback in depression. *Cognitive Therapy and Research*, *1*, 311–329.

Dennard, D. O., & Hokanson, J. E. (1986). Performance on two cognitive tasks by dysphoric and nondysphoric students. *Cognitive Therapy and Research*, *10*, 377–386.

*Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, *90*, 286–297.

DeRubeis, R. J., & Crits-Cristoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology*, *66*, 37–52.

DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology*, *58*, 862–869.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. C., Dunner, D. L., & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the treatment of adults with major depression. *Journal of Consulting and Clinical Psychology*, *74*, 658–670.

*Dobson, K. S., & Shaw, B. F. (1981). The effect of self-correction on cognitive distortions in depression. *Cognitive Therapy and Research*, *5*, 391–403.

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.

---

[6] References marked by an asterisk indicate studies included in the meta-analysis.

*Dobson, K. S., & Pusch, D. (1995). A test of the depressive realism hypothesis in clinically depressed subjects. *Cognitive Therapy and Research*, 19, 179–194.

*Dunn, B. D., Dalgleish, T., Lawrence, A. D., & Ogilvie, A. D. (2007). The accuracy of self-monitoring and its relationship to self-focused attention in dysphoria and clinical depression. *Journal of Abnormal Psychology*, 116, 1–15.

*Dunn, B. D., Stefanovitch, I., Buchan, K., Lawrence, A. D., & Dalgleish, T. (2009). A reduction in positive self-judgment bias is uniquely related to the anhedonic symptoms of depression. *Behaviour Research and Therapy*, 47, 374–381.

*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.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455–463.

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95, 89–98.

*Dykman, B. M., Abramson, L. Y., & Albright, J. S. (1991). Effects of ascending and descending patterns of success upon dysphoric and nondysphoric subjects' encoding, recall, and predictions of future success. *Cognitive Therapy and Research*, 15, 179–199.

*Dykman, B. M., Abramson, L. Y., Alloy, L. B., & Hartlage, S. (1989). Processing of ambiguous and unambiguous feedback by depressed and nondepressed college students: Schematic biases and their implications for depressive realism. *Journal of Personality and Social Psychology*, 56, 431–445.

*Dykman, B. M., Horowitz, L. M., Abramson, L. Y., & Usher, M. (1991). Schematic and situational determinants of depressed and nondepressed students' interpretation of feedback. *Journal of Abnormal Psychology*, 100, 45–55.

Evans, M. D., & Hollon, S. D. (1988). Patterns of personal and causal inference: Implications for the cognitive therapy of depression. In J. Garber, & M. E. P. Seligman (Eds.), *Human helplessness: Theory and application* (pp. 344–377). New York: Academic Press.

Field, A. P. (2003). The problem in using fixed-effects models of meta-analysis on real world data. *Understanding Statistics*, 2, 77–96.

*Finkel, C. B., Glass, C. R., & Merluzzi, T. V. (1982). Differential discrimination of self-referent statements by depressives and nondepressives. *Cognitive Therapy and Research*, 6, 173–183.

*Ford, C. E., & Neale, J. M. (1985). Learned helplessness and judgments of control. *Journal of Personality and Social Psychology*, 49, 1330–1336.

*Garber, J., & Hollon, S. D. (1980). Universal versus personal helplessness in depression: Belief in uncontrollability or incompetence. *Journal of Abnormal Psychology*, 89, 56–66.

Gilboa, E., & Gotlib, I. H. (1997). Cognitive biases and affect persistence in previously dysphoric and never-dysphoric individuals. *Cognition and Emotion*, 11, 517–538.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage Publications.

*Glass, D. C., McKnight, J. D., & Valdimarsdottir, H. (1993). Depression, burnout, and perceptions of control in hospital nurses. *Journal of Consulting and Clinical Psychology*, 61, 147–155.

Goddard, L., Dritschel, B., & Burton, A. (1996). Role of autobiographical memory in social problem solving and depression. *Journal of Abnormal Psychology*, 105, 609–616.

Golin, S., Terrel, F., Weitz, J., & Drost, P. L. (1979). The illusion of control among depressed patients. *Journal of Abnormal Psychology*, 88, 454–457.

Golin, S., Terrell, F., & Johnson, B. (1977). Depression and the illusion of control. *Journal of Abnormal Psychology*, 86, 440–442.

*Gotlib, I. H. (1981). Self-reinforcement and recall: Differential deficits in depressed and nondepressed psychiatric patients. *Journal of Abnormal Psychology*, 90, 521–530.

*Gotlib, I. H. (1982). Self-reinforcement and depression in interpersonal interaction: The role of performance level. *Journal of Abnormal Psychology*, 91, 3–13.

*Gotlib, I. H. (1983). Perception and recall of interpersonal feedback: Negative bias in depression. *Cognitive Therapy and Research*, 7, 399–412.

Gotlib, I. H., McLachlan, A. L., & Katz, A. N. (1988). Bias in visual attention in depressed and nondepressed individuals. *Cognition and Emotion*, 2, 185–200.

*Gotlib, I. H., & Meltzer, S. J. (1987). Depression and the perception of social skill in dyadic interaction. *Cognitive Therapy and Research*, 11, 41–54.

Greenberg, P. E., Stiglin, L. E., Finkelstein, S. N., & Berndt, E. R. (1993). The economic burden of depression in 1990. *The Journal of Clinical Psychiatry*, 54, 405–426.

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

*Hammen, C. L., & Krantz, S. (1976). Effect of success and failure on depressive cognitions. *Journal of Abnormal Psychology*, 85, 577–586.

*Hancock, J. A., Moffoot, A. P. R., & O'Carroll, R. E. (1996). "Depressive realism" assessed via confidence in decision-making. *Cognitive Neuropsychiatry*, 1, 213–220.

*Harkness, K. L., Sabbagh, M. A., Jacobson, J. A., Chowdrey, N. K., & Chen, T. (2005). Enhanced accuracy of mental state decoding in dysphoric college students. *Cognition and Emotion*, 19, 999–1025.

Hedges, L. V. (1981). Distribution theory for Glass' estimator of effect size and related estimators. *Journal of Educational Statistics*, 6, 107–128.

Hedges, L. V. (1994). Statistical considerations. In H. Cooper, & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 28–38). New York: Russell Sage.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.

Hedges, L. V., & Vevea, J. L. (1998). Fixed and random-effects models in meta-analysis. *Psychological Methods*, 3, 486–504.

Hoehn-Hyde, D., Schlottman, R. S., & Rush, A. J. (1982). Perception of social interactions in depressed psychiatric patients. *Journal of Consulting and Clinical Psychology*, 50, 209–212.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed-effects vs. random-effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8, 275–292.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology*, 64, 295–304.

*Javna, C. D. (1981). Depressed and nondepressed college students' interpretations of and memory for feedback about self and others. Unpublished doctoral dissertation, Ohio State University, Columbus, OH.

*Johnson, T. J., & DiLorenzo, T. M. (1998). Social information processing biases in depressed and nondepressed college students. *Journal of Social Behavior & Personality*, 13, 517–530.

*Johnson, J. E., Petzel, T. P., Hartney, L. M., & Morgan, R. A. (1983). Recall of importance ratings of completed and uncompleted tasks as a function of depression. *Cognitive Therapy and Research*, 7, 51–56.

*Kapci, E. G., & Cramer, D. (1998). The accuracy of dysphoric and nondepressed groups' predictions of life events. *Journal of Psychology: Interdisciplinary and Applied*, 132, 659–670.

*Kapci, E. G., & Cramer, D. (1999). Judgement of control revisited: Are the depressed realistic or pessimistic? *Counselling Psychology Quarterly*, 12, 95–105.

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 and Research*, 11, 289–299.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*, 289, 3095–3105.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 617–627.

*Klein, D. C. (1975). Reversal of performance deficits and perceptions of response-reinforcement independence associated with learned helplessness and depression. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia, PA.

*Krantz, S. E., & Gallagher-Thompson, D. (1990). Depression and information valence influence depressive cognition. *Cognitive Therapy and Research*, 14, 95–108.

Kuyken, W., & Brewin, C. R. (1995). Autobiographical memory functioning in depression and reports of early abuse. *Journal of Abnormal Psychology*, 104, 585–591.

Kuyken, W., & Dalgleish, T. (1995). Autobiographical memory and depression. *British Journal of Clinical Psychology*, 34, 89–92.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32, 311–328.

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

Lobitz, W. C., & Post, R. D. (1979). Parameters of self-reinforcement and depression. *Journal of Abnormal Psychology*, 88, 33–41.

*Loeb, A., Beck, A. T., & Diggory, J. C. (1971). Differential effects of success and failure on depressed and nondepressed patients. *The Journal of Nervous and Mental Disease*, 152, 106–114.

*Loewenstein, D. A., & Hokanson, J. E. (1986). The processing of social information by mildly and moderately dysphoric college students. *Cognitive Therapy and Research*, 10, 447–460.

Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavioral therapy? *Clinical Psychology Review*, 27, 173–187.

*Lovejoy, M. C. (1991). Maternal depression: Effects on social cognition and behavior in parent–child interactions. *Journal of Abnormal Child Psychology*, 19, 693–706.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional biases in emotional disorders. *Journal of Abnormal Psychology*, 95, 15–20.

*Margo, G. M., Greenberg, R. P., Fisher, S., & Dewan, M. (1993). A direct comparison of the defense mechanisms of nondepressed people and depressed psychiatric inpatients. *Comprehensive Psychiatry*, 34, 65–69.

*Martin, D. J., Abramson, L. Y., & Alloy, L. B. (1984). The illusion of control for self and others in depressed and nondepressed college students. *Journal of Personality and Social Psychology*, 46, 125–136.

McCabe, S. B., & Gotlib, I. H. (1995). Selective attention and clinical depression: Performance on a deployment-of-attention task. *Journal of Abnormal Psychology*, 104, 241–245.

McCabe, S. B., Gotlib, I. H., & Martin, R. A. (2000). Cognitive vulnerability to depression: Deployment-of-attention as a function of history of depression and current mood state. *Cognitive Therapy and Research*, 24, 427–444.

McCabe, S. B., & Toman, P. E. (2000). Stimulus exposure duration in a deployment-of-attention task: Effects on dysphoric, recently dysphoric, and nondysphoric individuals. *Cognition and Emotion*, 14, 125–142.

*McKendree-Smith, N., & Scogin, F. (2000). Depressive realism: Effects of depression severity and interpretation time. *Journal of Clinical Psychology*, 56, 1601–1608.

*McNamara, K., & Hackett, G. (1986). Gender, sex-type and cognitive distortion: Self-perceptions of social competence among mild depressives. *Social Behavior and Personality*, 14, 113–121.

*Mikulincer, M., Gerber, H., & Weisenberg, M. (1990). Judgment of control and depression: The role of self-esteem threat and self-focused attention. *Cognitive Therapy and Research*, 14, 589–608.

*Miller, W. R., & Seligman, M. E. P. (1973). Depression and the perception of reinforcement. *Journal of Abnormal Psychology*, 82, 62–73.

*Miller, W. R., & Seligman, M. E. P. (1976). Learned helplessness, depression, and the perception of reinforcement. *Behaviour Research and Therapy*, 14, 7–17.

*Miller, W. R., Seligman, M. E. P., & Kurlander, H. M. (1975). Learned helplessness, depression, and anxiety. *The Journal of Nervous and Mental Disease*, 161, 347–357.

Moore, M. T., & Fresco, D. M. (2007). Depressive realism and attributional style: Implications for individuals at risk for depression. *Behavior Therapy*, 38, 144–154.

*Moretti, M. M. (1985). Processing positive and negative stimuli in depression. Unpublished doctoral dissertation, Simon Fraser University, Vancouver, BC, Canada.

*Msetfi, R. M., Murphy, R. A., & Simpson, J. (2007). Depressive realism and the effect of inter-trial interval on judgments of zero, positive and negative contingencies. *Quarterly Journal of Experimental Psychology*, *60*, 461–481.

*Msetfi, R. M., Murphy, R. A., Simpson, J., & Kornbrot, D. E. (2005). Depressive realism and outcome density bias in contingency judgments: The effect of the context and intertrial interval. *Journal of Experimental Psychology: General*, *134*, 10–22.

Musson, R. F., & Alloy, L. B. (1987). Depression, self-consciousness, and judgments of control: A test of the self-focused attention hypothesis. Unpublished manuscript, Northwestern University, Evanston, IL.

*Nelson, R. E., & Craighead, W. E. (1977). Selective recall of positive and negative feedback, self-control behaviors and depression. *Journal of Abnormal Psychology*, *86*, 379–388.

*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.

*Presson, P. K., & Benassi, V. A. (2003). Are depressive symptoms positively or negatively associated with the illusion of control? *Social Behavior and Personality*, *31*, 483–495.

Puffet, A., Jehin-Marchot, D., Timsit-Berthier, M., & Timsit, M. (1991). Autobiographical memory and major depressive states. *European Psychiatry*, *6*, 141–145.

*Pusch, D., Dobson, K. S., Ardo, K., & Murphy, T. (1998). The relationships between sociotropic and autonomous personality styles and depressive realism in dysphoric and nondysphoric university students. *Canadian Journal of Behavioural Science*, *30*, 253–265.

*Pyszczynski, T., Holt, K., & Greenberg, J. (1987). Depression, self-focused attention, and expectancies for positive and negative future life events for self and others. *Journal of Personality and Social Psychology*, *52*, 994–1001.

Raudenbush, S. W. (1994). Analyzing effect sizes: Random effects models. In H. Cooper, & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 295–315). New York: Russell Sage.

*Rosenfarb, I. S., Burker, E. J., Morris, S. A., & Cush, D. T. (1993). Effects of changing contingencies on the behavior of depressed and nondepressed individuals. *Journal of Abnormal Psychology*, *102*, 642–646.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park, CA: Sage.

*Roth, D., & Rehm, L. P. (1980). Relationships among self-monitoring processes. Memory, and depression. *Cognitive Therapy and Research*, *4*, 149–157.

Rozensky, R. H., Rehm, L. P., Pry, G., & Roth, D. (1977). Depression and self-reinforcement behavior in hospitalized patients. *Journal of Behavior Therapy and Experimental Psychiatry*, *8*, 35–38.

Ruehlman, L. S., West, S. G., & Pasahow, R. J. (1985). Depression and evaluative schemata. *Journal of Personality*, *53*, 46–92.

*Sacco, W. P., & Hokanson, J. E. (1978). Expectations of success and anagram performance of depressives in a public and private setting. *Journal of Abnormal Psychology*, *87*, 122–130.

*Sacco, W. P., & Hokanson, J. E. (1982). Depression and self-reinforcement in a public and private setting. *Journal of Personality and Social Psychology*, *42*, 377–385.

Schmidt, F. L., Oh, I., & Hayes, T. L. (2009). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*, *62*, 97–128.

Siegel, S. J., & Alloy, L. B. (1990). Interpersonal perceptions and consequences of depressive-significant other relationships: A naturalistic study of college roommates. *Journal of Abnormal Psychology*, *99*, 361–373.

*Stone, A. L., & Glass, C. R. (1986). Cognitive distortion of social feedback in depression. *Journal of Social and Clinical Psychology*, *4*, 179–188.

*Stone, E. R., Dodrill, C. L., & Johnson, N. (2004). Depressive cognition: A test of the depressive realism versus negativity using general knowledge questions. *The Journal of Psychology*, *135*, 583–602.

*Strack, S., & Coyne, J. C. (1983). Social confirmation of dysphoria: Shared and private reactions. *Journal of Personality and Social Psychology*, *44*, 798–806.

*Strunk, D. R., & Adler, A. D. (2009). Cognitive biases in three predictive tasks: A test of the cognitive model of depression. *Behaviour Research and Therapy*, *47*, 34–40.

*Strunk, D. R., Lopez, H., & DeRubeis, R. J. (2006). Depressive symptoms are associated with unrealistic negative predictions of future life events. *Behaviour Research and Therapy*, *44*, 875–896.

Tennen, H., & Herzberger, S. (1987). Depression, self-esteem, and the absence of self-protective attributional biases. *Journal of Personality and Social Psychology*, *52*, 72–80.

*Vazquez, C. V. (1987). Judgment of contingency: Cognitive biases in depressed and nondepressed subjects. *Journal of Personality and Social Psychology*, *52*, 419–431.

*Vestre, N. D., & Caulfield, B. P. (1986). Perception of neutral personality descriptors by depressed and nondepressed subjects. *Cognitive Therapy and Research*, *10*, 31–36.

Wener, A. E., & Rehm, L. P. (1975). Depressive affect: A test of behavioral hypothesis. *Journal of Abnormal Psychology*, *84*, 221–227.

*Wenzlaff, R. M. (1984). Depression and judgments of personal feedback: Accuracy at the expense of self-enhancement? Unpublished doctoral dissertation, University of Texas, Austin.

Wenzlaff, R. M., & Berman, J. S. (1985, August). Judgmental accuracy in depression. *Paper presented at the meeting of the American Psychological Association, Los Angeles.*

*Whitton, S. W., Larson, J. J., & Hauser, S. T. (2008). Depressive symptoms and bias in perceived social competence among young adults. *Journal of Clinical Psychology*, *64*, 791–805.