Handbook of Implicit Social Cognition

Chapter: Clinical applications of implicit social cognition theories and methods

Bethany A. Teachman, Meghan W. Cody, and Elise M. Clerkin

University of Virginia
It is exciting to see the implicit social cognition and psychopathology research traditions come together, because the two fields have often operated independently. The historic neglect of one another’s findings has been unfortunate, because in many ways the two fields are vitally linked. To understand abnormal, maladaptive ways of processing information about the self and others, it is critical to understand how this occurs in healthy populations and is influenced by situational factors. Analogously, to understand the influence of the social environment on cognitive processing, it is essential to recognize the full range of individual differences and how this intra- and interpersonal variability and the environment interact. Moreover, the influences of emotions, attitudes, and beliefs on behavior are at the core of both fields, so integrating their findings holds promise to move each field forward. We are increasingly able to ask important questions that would not be possible without the influence of one another’s methods and perspectives about how human beings make sense of their social world.

In this chapter, we focus primarily on indirect measures of disorder-relevant attitudes and beliefs\(^1\) that capture some aspect of implicit social cognitive processing (i.e., reflecting one of the features of automaticity outlined by Bargh, 1994: uncontrollable, unintentional, efficient, or unconscious). Clinical applications of implicit association and priming measures are emphasized given the central role of these paradigms in current social cognition research. At the same time, because much of the work in psychopathology has used more traditional measures of information processing, such as indicators of attentional interference like the emotional Stroop (labeled e-Stroop to differentiate it from the classic Stroop) and dot probe tasks, we will also briefly summarize evidence from these tasks.

Evidence from the selected paradigms has been organized into broad disorder categories, reflecting the most common problem areas and those with the strongest relevant empirical literatures: anxiety, mood, eating, body dysmorphic, psychotic, and personality disorders (substance abuse research is covered in Wiers et al., this volume). Within each disorder category, we examine evidence for group differences between clinical and nonclinical samples, predictive validity, vulnerability to psychopathology or causal links between biased processing and changes in symptoms, and relations between task performance and treatment outcome. Finally, we consider future research directions needed to better understand implicit social cognitive processes in clinical populations. Emphasis has been placed on

---

1 There is a small but growing literature examining indirect measures of stereotypes tied to persons with mental illness that is also relevant, which is discussed in Amodio and Mendoza (this volume).
diagnosed samples, but reports of analogue samples (groups that are high in symptoms of a given disorder, but have not been formally diagnosed) are also included. This inclusion was based on recognition of the continuous, dimensional nature of many areas of psychopathology (see Brown & Barlow, 2005) and to provide more comprehensive coverage of the current state of the literature. Also, we focus on the adult literature, but have noted a couple interesting examples from the child literature when they are relevant to disorder vulnerability.

Anxiety Disorders

Anxiety disorders are characterized by excessive fear, worry, panic or anxiety that is sufficiently severe that it impairs a person’s functioning. Based on the current version of the Diagnostic and Statistical Manual (DSM IV-TR: American Psychiatric Association, 2000), there are seven different anxiety disorders: 1) panic disorder, characterized by recurrent panic attacks that involve the rapid onset of multiple physical symptoms (such as racing heart and shortness of breath) and fear of future attacks; 2) agoraphobia, where a person avoids situations where anxiety or panic symptoms might occur; 3) social anxiety disorder, which involves fear of negative evaluation by others and consequent avoidance of interpersonal and/or performance situations; 4) specific phobias, exemplified by excessive fear toward a particular target, such as animals, certain natural environments (e.g., heights), closed spaces, etc.; 5) obsessive-compulsive disorder (OCD), characterized by repeated unwanted intrusive thoughts or images (obsessions) and/or ritualized behaviors designed to reduce distress (compulsions; e.g., washing one’s hands excessively); 6) generalized anxiety disorder (GAD), which involves extreme worry and catastrophic thinking about a wide range of situations; and 7) posttraumatic stress disorder (PTSD), in which exposure to a trauma is followed by re-experiencing of the event in some way (e.g., in dreams or flashbacks), avoidance of cues associated with the trauma, and increased arousal. As evident from this list, the target of the fear varies across the different anxiety disorders, but in all cases the fear or worry is excessive relative to what others would experience in that situation. In addition, individuals with these disorders usually go to great lengths to avoid their feared situation, which adds considerably to the associated impairment in occupational and social functioning.

There is a rich history of theoretical and empirical work noting the biased nature of automatic processing in anxiety disorders (e.g., Beck & Clark, 1997; Mathews & MacLeod, 2005; McNally, 1995;
Mogg & Bradley, 2005; Williams, Watts, MacLeod, & Mathews, 1997). In Beck’s early writings he pointed to the role of “automatic thoughts” in anxiety problems, intrusive thoughts whose content involved danger or harm (Beck & Emery with Greenberg, 1985). While these thoughts may be considered automatic insofar as they are likely somewhat involuntary, they are also the conscious output of elaborative processes. In more recent writings, Beck and colleagues outline a model that more explicitly delineates periods where the relative balance between automatic and strategic processing will vary. In particular, Beck and Clark (1997) propose a schema-based information processing model of anxiety “that involves: (a) the initial registration of a threat stimulus (orienting mode); (b) the activation of a primal threat mode; and (c) the secondary activation of more elaborative and reflective modes of thinking” (p. 49). In this model, automatic processing is expected to be active throughout all three stages, whereas the role for strategic processing increases over time.

Other researchers have focused on the presence of different features of automaticity in anxiety, rather than on changes over time. In particular, in a seminal paper, McNally (1995) reviewed evidence for capacity-free, unconscious, and involuntary automatic processing in anxiety, concluding that the support was strongest for involuntary processing and relatively weak for capacity-free or unconscious processing: “it is the inability of the patient to terminate fear-generating processing once it starts that is the hallmark of pathological anxiety” (p. 752). This conclusion was based on the available paradigms at that time to assess the selective processing of potentially threatening information. With the advent of a broader range of paradigms and tighter controls within tasks (e.g., for assessing unconscious processing), we can now evaluate these claims more fully.

**Group Differences and Predictive Validity**

**Implicit Association Tasks**

Known groups differences. Over the past decade, there has been a growing interest in examining whether persons with anxiety difficulties differ from non-anxious individuals in their implicit associations, either related to the self or to the feared object. While the majority of demonstrations have been conducted with the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), related measures, such as the Go/No-go Association Task (GNAT; Nosek & Banaji, 2001) and Extrinsic Affective Simon Task (EAST; De Houwer, 2003), have also been used. (See description of this class of measures
It is clear at this stage that at least some anxiety problems are distinguished by automatic fear associations in memory, though more work with diagnosed samples across the range of anxiety disorders is sorely needed.

A great deal of this work has occurred in the area of specific animal fears and phobias, particularly spider fear, likely because these samples are readily accessible and the fear target is clearly specified. In an early demonstration that the IAT could distinguish between persons with fears of either snakes or spiders, Teachman, Gregg, and Woody (2001) found an opposite pattern of implicit associations for the different fear groups on four IAT tasks that involved classification of pictures of snakes and spiders with descriptive words indicating valence (bad/good), fear, danger, or disgust. In addition, the differential association of the animals with being afraid (versus unafraid) was significant even after controlling for the IAT effect from the implicit bad/good evaluation, suggesting that fear associations are specific not only to the particular fear target (i.e., snakes versus spiders) but also capture fear group differences beyond simple liking. Since this initial demonstration, a number of other researchers have also demonstrated specific spider fear associations using the IAT (e.g., Ellwart, Rinck, & Becker, 2006; Huijding & de Jong, 2005).

Similar associations have been evident on related tasks that do not require a specific, relative comparison target category as the IAT does. For instance, Teachman (2007) found stronger implicit spider fear associations on the GNAT among spider fearful, relative to non-fearful, individuals. On the EAST, Huijding and de Jong (2005) differentiated among high and low spider fear groups in their automatic responses to spider pictures, and Ellwart, Becker, and Rinck (2005) found similar group differences with spider-relevant words, but only when the concept of spiders had been primed. Notably, there have been failures to replicate these spider fear group differences (e.g., de Jong, van den Hout, Rietbroek, & Huijding, 2003), but the overall pattern of implicit fear associations is clear.

Importantly, the evidence for implicit associations is not limited to spider fear. In the area of specific fears and phobias more generally, there is evidence based on the IAT for greater implicit height fear associations among people high (versus low) in acrophobia symptoms (Teachman, Stefanucci, Clerkin, Cody, & Proffitt, 2008), and for greater disgust associations with blood and injection images for people high (versus low) in blood-injury-injection symptoms (Saporito, Heiman, Unger, & Teachman,
There is also evidence for group differences on the IAT in associating “panic” versus “calm” with the self among people high (versus low) in anxiety sensitivity (a vulnerability marker for anxiety problems characterized by fear of anxiety symptoms; Teachman, 2005), and among individuals diagnosed with panic disorder, relative to healthy control participants (Teachman, Smith-Janik, & Saporito, 2007). There is less clear evidence in the area of PTSD, though see Engelhard, Huijding, van den Hout, and de Jong (2007). Finally, in the area of social phobia, both de Jong (2002) and Tanner, Stopa, and De Houwer (2006) found relatively less positive implicit self-esteem among women who were high, relative to low, in social anxiety symptoms. Similarly, de Jong, Pasman, Kindt, and van den Hout (2001) found a tendency for women high in social anxiety to associate social situations with more negative outcomes, relative to women low in social anxiety (though Teachman & Allen, 2007, did not find a relationship between implicit and explicit measures tied to social anxiety in an unselected adolescent sample). Thus, although far more work is needed with diagnosed samples, there are now considerable data suggesting the expected group differences in disorder-relevant implicit associations across numerous anxiety difficulties.

Predictive validity. The evidence for predictive validity is somewhat less extensive but also promising. There are many examples of implicit associations predicting avoidance behavior and distress when participants are asked to approach their feared object (e.g., avoidance of a spider: Ellwart et al., 2006; reactions to a balcony and ladder in height fear: Teachman et al., 2008), including support for incremental validity. For instance, implicit spider fear associations predicted degree of approach to a live spider above and beyond the variance explained by self-report measures (e.g., Teachman, 2007; Teachman & Woody, 2003; see also prediction of anxiety during a stressful speech task in an unselected sample: Egloff & Schmukle, 2002). In panic disorder, relatively lower implicit associations of the self as calm predicted a latent factor reflecting a variety of panic disorder symptoms, including panic frequency, agoraphobic avoidance, and distress during a task that elicited lightheadedness (a common panic symptom; Teachman et al., 2007). Interestingly, in this study, the implicit associations did not significantly correlate with other measures of information processing biases (e.g., the e-Stroop task), suggesting good prediction of symptoms, but some independence among different aspects of biased cognitive processing. Further, it appears that implicit and explicit fear measures may at times have differential predictive validity, with some hints that implicit associations may be more strongly associated with relatively more
automatic, versus strategic, measures tied to anxiety-relevant markers (though this is not a consistent finding). Huijding and de Jong (2006), for example, found that implicit spider fear associations predicted eye blink startle response (often interpreted as a measure of avoidance motivation; Lang, Bradley, & Cuthbert, 1990) more strongly than a self-report measure, which in turn was a better predictor of more strategic avoidance behavior.

**Priming Tasks**

Unlike the extensive work in the social cognition field using priming tasks, there are fewer examples of semantic or affective priming paradigms being applied with clinically anxious populations, so drawing general conclusions about priming effects is somewhat premature (though see intriguing evidence that trait anxiety may moderate affective and conceptual priming effects in nonclinical samples; e.g., Berner & Maier, 2004; Fox, 1994; Li, Paller, & Zinbarg, 2008; Maier, Berner, Hau, & Pekrun, 2007). This is unfortunate because the designs often lend themselves well to demonstrating uncontrollable processing, arguably the strongest feature of automaticity in anxiety disorders (McNally, 1995).

Negative priming paradigms seem particularly useful because they have the potential to reflect both uncontrollable and unintentional processing, depending on the particular task instructions and design. This class of priming paradigms has been used most frequently with OCD samples to examine general cognitive processing differences and learning deficits in this population (e.g., Enright & Beech, 1990; McNally, Wilhelm, Buhlmann, & Shin, 2001), such as difficulties with selective inhibition. However, it has not typically been used to examine reactions to disorder-relevant content. An interesting exception comes from a study by Amir, Cobb and Morrison (2008) in which they used a modified negative priming paradigm that included idio graphically selected words for individuals with OCD (e.g., “contaminated” for a person with washing concerns) and yoked, non-anxious control participants. They found opposite priming effects for neutral versus threat words in the OCD sample, which they interpreted to suggest that individuals with OCD have difficulty learning contingencies when information is personally relevant to their fears.

Further supporting the importance of idio graphically matching stimuli so that it is personally relevant, Schneider and Schulte (2007) found that individuals with panic disorder demonstrated stronger semantic priming effects than did a nonclinical control group when naming catastrophic target words (e.g.,
“infarction”)) immediately following aurally presented prime sentences that described the perception of anxiety symptoms. Notably, there was no group difference when looking across identical stimuli for all participants (see also Schniering & Rapee, 1997); the effect was only evident when priming effects were calculated for idiographically selected stimuli. Thus, while evidence is sparse at this stage, early findings suggest it is important to take into account the heterogeneous nature of threat stimuli when examining priming effects in anxious samples.

Also, there is some evidence that small variations in stimuli exposure can influence the expression of priming results. McNally, Hornig, Otto and Pollack (1997) found evidence for priming effects for threat relative to positive words for persons diagnosed with panic disorder (but not for a non-anxious control group), but the effect was only evident for targets presented in the bottom (versus top) of the screen. Moreover, timing of stimuli presentation is clearly important, with some studies only seeing effects at relatively longer stimulus-onset-asynchronies (SOAs; e.g., 2000 ms), suggesting that relatively more controlled processing was likely influencing results (see also Clark et al., 1988). Similarly, Dalgleish, Cameron, Power, and Bond (1995) found faster endorsement of negative self-descriptive adjectives following negatively valenced primes for a GAD sample at 2000 ms (versus at 250 ms) SOA. Notably, for other information processing paradigms, such as the dot probe task (described below), results are often more robust with anxious samples at shorter SOAs, so the optimal timing of stimuli presentation remains an important but unresolved question across paradigms.

We suspect that these highly specific requirements for stimuli presentation and the heterogeneity of concerns among anxious samples may partly explain some prior null findings (e.g., Kindt & Brosschot, 1998) in the priming literature. For example, Bradley, Mogg, and Williams (1995) did not find an effect of priming on a lexical decision task among a mixed clinically anxious group. Taken together, we see priming paradigms as potentially useful for examining automatic processes in anxiety, but this work is still in its infancy. One possible reason the paradigms have been used less frequently concerns the oft-cited problem of low reliability of priming measures, which can make them difficult to use for individual differences and repeated measures designs that are common for psychopathology and treatment studies.

**Attention Bias Tasks**
Although measures of attentional bias, such as the e-Stroop and dot probe task, are used less frequently in the implicit social cognition field, these tasks can be effective at capturing some features of automatic processing, such as uncontrollability and lack of conscious awareness (when stimuli are presented subliminally); thus, we briefly summarize some of the key findings from these tasks.

To investigate unconscious processing, the subliminal e-Stroop and modified dot probe tasks have been used across a range of anxious populations. In these paradigms, latency to complete a task, such as ink color naming or detecting the location of a probe, is compared when fear-relevant versus fear-irrelevant stimuli are presented very briefly and then backward masked. Findings for both the subliminal e-Stroop and dot probe paradigms have been somewhat inconsistent (regarding whether group differences are observed in interference effects following presentation of the fear-relevant stimuli).

There is little doubt that unconscious processing of fear stimuli can occur given that there are many positive findings on the subliminal e-Stroop (e.g., panic disorder with agoraphobia: Lundh, Wikstrom, Westerlund, & Öst, 1999; GAD: Bradley, Mogg, Millar, & White, 1995; Mogg, Bradley, Williams, & Mathews, 1993; PTSD: Harvey, Bryant, & Rapee, 1996; specific phobias: van den Hout, Tenney, Huygens, & de Jong, 1997), and to a lesser extent on the subliminal version of the dot probe (e.g., Mogg, Bradley, & Williams, 1995). However, there are also many null findings, suggesting a number of likely moderators of these effects. For the subliminal dot probe, results appear to be stronger for clinical, relative to subthreshold analogue samples, and stronger when samples are homogenous with respect to the specific anxiety disorder being examined; however, low reliability of the task has been a serious constraint (see review by Schmukle, 2005). For the subliminal e-Stroop, methodological differences across studies may explain some of the null findings, including variable approaches to masking, stimuli presentation durations, presence of a stressor during the task, and small sample sizes (see also Williams, Mathews, & MacLeod, 1996, for review of the likely mechanisms underlying e-Stroop interference). Also, the evidence appears to be more robust for some anxiety disorders over others (e.g., results are quite mixed in specific phobias; Wenzel & Holt, 1999), though methodological variations across studies confound a simple disorder-specific interpretation.

More consistent evidence for phobic group differences has been observed using paradigms that present masked stimuli and then examine how quickly fear responses can be conditioned or extinguished.
to fear-relevant stimuli (see review by Öhman & Mineka, 2001). These conditioning/extinction tasks do not necessarily investigate attentional processes directly, but do support claims that fearful responses can be activated by unconscious perceptual analysis of fear-relevant stimuli.

Supraliminal versions of the e-Stroop and dot probe tasks have been used extensively with anxiety disordered populations (see also related measures such as the Posner task; e.g., Amir, Elias, Klumpp, & Przeworski, 2003). These tasks seem to capture uncontrollable processing well, and, in some cases, may also reflect unintentional processing of emotional information. There are a great many studies using the e-Stroop paradigm, and despite a number of null results (e.g., Kampman, Keijsers, Verbraak, Näring, Hoogduin, 2002; Moritz et al., 2004), there is a plethora of examples indicating that anxious individuals have selective interference when processing fear-relevant stimuli, even though task performance does not require semantic processing of the stimuli. We list just a couple examples from the various anxiety disorders here, and encourage the interested reader to see recent helpful reviews and meta-analyses by Mogg and Bradley (2005), Phaf and Kan (2007), and Ehrenreich and Gross (2002, for a review of the child anxiety literature).

Supraliminal e-Stroop effects have been observed in each of the anxiety disorders (e.g., specific phobia: Thorpe & Salkovskis, 1997; Wikström, Lundh, Westerlund, & Högman, 2004; social phobia: Becker, Rinck, Margraf, & Roth, 2001; Spector, Pecknold, & Libman, 2003; GAD: Becker et al., 2001; Mogg, Bradley, Millar, & White, 1995; PTSD: McNally, Kaspi, Riemann & Zeitlin, 1990; Paunovic, Lundh, & Öst, 2002; Vythilingam et al., 2007; OCD: Foa, Iliai, McCarthy, Shoyer & Murdock, 1993; Hartston & Swerdlow, 1999; panic disorder with and without agoraphobia: Lundh et al., 1999; McNally, Riemann & Kim, 1990). At the same time, despite the many examples of significant e-Stroop effects, a number of factors seem to influence the likelihood of observing strong effects: the specific anxiety disorder examined (e.g., stronger effects for PTSD, and somewhat less robust evidence for panic disorder and OCD); stimuli characteristics (e.g., interference effects are enhanced when the task utilizes personally relevant negative stimuli; see Williams et al., 1996); and comorbidity (e.g., effects appear to diminish when anxiety is comorbid for depression; see Bradley et al., 1995). Based on their meta-analysis, Phaf and Kan (2007) argue that e-Stroop effects are more robust for the supra- versus subliminal version and when words are presented in blocked format, which they interpret as evidence that the interference effect is driven by a
relatively slow (versus fast) process whereby the anxious individual has difficulty disengaging from threat stimuli (suggesting uncontrollability).

The dot probe task has been used less extensively with diagnosed anxious samples. However, there are still numerous examples of enhanced identification of the probe following threat-relevant (relative to control) cues, suggesting attention was already oriented towards the threat information. For instance, group differences have been observed in GAD (e.g., Bradley, Mogg, White, Groom, & de Bono, 1999; Mogg, Mathews, & Eysenck, 1992), panic disorder (e.g., Kroeze & van den Hout, 2000; Horenstein & Segui, 1997), social phobia (e.g., Musa, Lépine, Clark, Mansell, & Ehlers, 2003, if the socially phobic participants did not also have comorbid depression), and in specific fear and phobic samples (e.g., blood-injury: Mogg, Bradley, Miles, & Dixon, 2004; spiders: Mogg & Bradley, 2006, when stimuli were presented at 200 ms). Again, as with the e-Stroop results, there are also a number of null results (e.g., PTSD: Elsesser, Sartory, & Tackenberg, 2004; specific phobias: Elsesser, Heuschen, Pundt, & Sartory, 2006; Wenzel & Holt, 1999). Factors that seem to influence the likelihood of observing group differences include having relatively large sample sizes to capture smaller effect sizes, and the ideal timing of stimuli presentation appears to vary across disorders. For instance, Mogg and Bradley (2006) suggest stronger results in specific phobias when stimuli are presented for shorter durations (e.g., 200 ms), but results are found more consistently in GAD at longer durations (e.g., 500 ms). This difference points to likely variation across the disorders in how rapidly attention is first oriented toward threat stimuli and the rate at which disengagement of attention occurs. Along these lines, there are intriguing signs that under some conditions a bias away from threat-relevant information may be evident on the dot probe task (e.g., social phobia: Chen, Ehlers, Clark, & Mansell, 2002). These findings raise questions about when attentional orienting versus disengagement (which may reflect avoidance2) will occur in anxious populations.

Clearly, learning more about the time course of these effects is important, but results across tasks suggest that seemingly uncontrollable processing of emotional information influences subsequent task performance, and there are hints supporting these tasks’ predictive validity (e.g., Elsesser, 2006, found that heart rate in reaction to fear-relevant stimuli was associated with attentional biases).

---

2 This idea is popular among ‘vigilance-avoidance’ models that suggest anxious individuals rapidly orient to threat stimuli but are then motivated to avoid further processing of the noxious stimuli (see Amir, Foa, & Coles, 1998).
This conclusion is further strengthened by a class of tasks known as visual search tasks that are premised upon the idea that speed of detecting a target cue embedded in a matrix of other cues will vary as a function of whether the target and/or distracter cues are fear-relevant. These tasks again point to an attentional bias in anxious samples, and provide some evidence of uncontrollable processing (particularly under conditions when the distracter items are threat-relevant). The primary evidence comes from phobic groups, including spider fear and phobia (Miltner, Krieschel, Hecht, Trippe, & Weiss, 2004; Rinck, Reinecke, Ellwart, Heuer, & Becker, 2005) and women with social phobia (Rinck & Becker, 2005), though see recent mixed evidence from a sample with PTSD (Pineles, Shipherd, Welch, & Yovel, 2007). In general, findings suggest that even when task performance would be enhanced by counteracting the activation of emotional information (e.g., ignoring spider distracter cues), interference effects indicate that anxious, relative to non-anxious, participants have greater difficulty controlling this process. Further reinforcing this conclusion is emerging evidence using a recently developed task, the Approach-Avoidance Task (AAT; e.g., Chen & Bargh, 1999), which also seems to capture attentional interference. This task involves comparing time to approach versus avoid different types of stimuli (typically by pushing or pulling a joystick) when the approach/avoid instructions are based on a fear-irrelevant feature (e.g., the shape of the picture stimulus) and the content of the stimuli should theoretically be ignored. Greater avoidance tendencies toward threat-relevant stimuli have been observed among socially anxious (Heuer, Rinck & Becker, 2007) and spider fearful (e.g., Rinck & Becker, 2007) individuals, where the AAT also predicted actual approach behavior to a live spider.

We have only reviewed a sample of the tasks that have been used to demonstrate automatic processes in anxious samples, highlighting those that have the closest ties to the implicit social cognition literature and/or strongest empirical base in the psychopathology literature. Notwithstanding the brevity of this review, the convergent evidence supports the utility of these paradigms to differentiate anxious from non-anxious samples. Further, effects on these tasks show important relationships to a variety of symptom and behavioral measures. Limitations of the existing literature highlight the need for more studies with diagnosed samples and increased understanding of the factors that moderate the strength of the observed effects (especially given the many null findings across paradigms, and sensitivity of the
tasks to minor procedural variations). Also, the evidence cited thus far has mostly been cross-sectional; we consider the smaller body of evidence speaking to temporal prediction and causal relationships next.

**Ties to Vulnerability and Causal Links**

Most of the work examining vulnerability and causal links between measures relevant to implicit social cognition and anxiety disorders has occurred with attentional bias paradigms, such as the e-Stroop and dot probe. For instance, greater interference effects on the e-Stroop task predicted more intense emotional responding to a common panic trigger (a CO$_2$ challenge; Nay, Thorpe, Roberson-Nay, Hecker, & Sigmon, 2004), and a more negative emotional response following a stressful medical diagnostic procedure (MacLeod & Hagan, 1992). Similarly, Egloff, Wilhelm, Neubauer, Mauss, and Gross (2002) found that performance on a dot probe task predicted subsequent cardiovascular reactivity (a common physiological marker of fear) one week later during a speech stressor task in an unselected sample.

Less work tied to vulnerability has been conducted with implicit association or priming paradigms, likely because these paradigms are relative newcomers in the psychopathology field. Thus, while some interesting studies are underway - for instance, to examine implicit associations as predictors of future changes in anxiety symptoms - results from this work are not yet available. One early null finding is that an IAT administered to Dutch soldiers following deployment to Iraq did not predict later development of PTSD (Engelhard et al., 2007). However, it is simply too early at this stage to know whether these measures will serve as markers of vulnerability. A positive sign in this regard was the finding that implicit panic associations were elevated among persons with high anxiety sensitivity even when this group reported no experience (yet) with clinical panic (Teachman, 2005). This suggests that the implicit associations were tied to a vulnerability marker for anxiety disorders (i.e., anxiety sensitivity), rather than occurring as a consequence of experiencing panic. Also, manipulations designed to shift beliefs that are thought to constitute vulnerability markers for OCD (specifically, shifting interpretations that intrusive thoughts are an important window into a person’s morals and character) also shifted related implicit associations about intrusive thoughts and implicit self-evaluations, especially for those individuals with pre-existing obsessional beliefs that likely made them more susceptible to the manipulation (Teachman & Clerkin, 2007; Teachman, Woody, & Magee, 2006). These findings again suggest an association
between implicit associations and markers for anxiety vulnerability in non-clinical samples, but
prospective longitudinal investigations are clearly needed.

Work examining a direct causal relationship between changes in measures relevant to implicit
social cognition and consequent change in anxiety is at very early stages, and again the majority of this
work has occurred with attentional bias paradigms, particularly the modified dot probe. Recent research
has shown that by training participants to expect that a probe will appear more consistently following non-
threat relevant stimuli (rather than equally following threat and non-threat stimuli), participants can learn
to attend to neutral stimuli. Importantly, this training can influence emotional responding to subsequent
stressors (e.g., MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Mathews & MacLeod, 2002),
including the finding that attentional training helped Singaporean high school students adjust to
emigration to Australia (See, MacLeod, & Bridle, 2009). Initial applications of this training program with
anxious samples have also been promising. Amir, Weber, Beard, Bomyea, and Taylor (2008) found that
a single session of training with a socially anxious sample resulted in lower self-reported anxiety following
a speech and higher ratings of speech performance by blind raters (compared to socially anxious
participants in an attentional control condition). Similarly, repeated attentional training over a seven-day
period with a socially anxious sample led to reduced scores on the Social Interaction Anxiety Scale,
compared to a control group (though not all measures of social anxiety were reduced, suggesting
circumscribed training effects; Li, Tan, Qian, & Liu, 2008). Analogously, an eight-session attentional
training program with persons with GAD resulted in lower anxiety on both self-report and interviewer
measures, relative to a control group (Amir, Beard, Burns, & Bomyea, 2009). Although this work is quite
novel, the accumulating evidence provides exciting support for a causal connection between less biased
processing of threat information and anxiety reduction.

We are aware of only one study to date that has directly tried to train implicit associations with an
anxious sample. Clerkin and Teachman (2009) used a classical conditioning paradigm to train socially
anxious participants to associate their speech performance with positive evaluations from others (through
repeated pairings of self-relevant stimuli and faces indicating positive social feedback). This positive
training, relative to two control conditions, led participants to speak longer during a subsequent speech
challenge (though not to report less anxiety) and to display less implicit rejection associations following
These preliminary findings suggest that implicit association training may also be causally related to reductions in anxiety, akin to attentional training, but further work is needed with diagnosed samples, more intensive training designs, and longer follow-up periods to determine the durability of training.

**Relation to Treatment**

There is a small but growing body of evidence suggesting that certain implicit associations are altered following cognitive behavior therapy for anxiety disorders (e.g., fear and danger associations in spider phobia: Teachman & Woody, 2003; self with panic associations in panic disorder: Teachman, Marker & Smith-Janik, 2008), though this has not been a consistent finding (e.g., Huijding and de Jong did not find treatment effects following exposure therapy for spider phobia beyond changes seemingly due to practice effects using either an IAT, 2007a, or an EAST, 2007b). The reason for the discrepant results across the spider phobia studies is not entirely clear, though one explanation may be that the Huijding and de Jong treatment occurred within one session, whereas the Teachman and Woody treatment was spaced out over three sessions. The extra time may have allowed more opportunity for consolidation of the new associative learning from treatment, resulting in a more substantive effect on the measures of implicit associations. However, the Teachman and Woody (2003) study did not include an untreated anxious control group, so it is difficult to directly evaluate this hypothesis without further data.

Importantly, there is also some evidence that change in implicit associations may actually predict treatment response (as opposed to simply co-occurring with change in treatment). Teachman, Marker, and Smith-Janik (2008) found that the trajectory of change on a measure of implicit panic associations predicted future change in panic symptoms over the course of therapy for panic disorder.

There are few if any evaluations of change in priming measures following treatment for anxiety problems, but there are some nice demonstrations using attentional bias paradigms. There is evidence, for example, that e-Stroop scores not only change over the course of treatment (e.g., van den Hout et al., 1997), but e-Stroop effects appear to be preferentially reduced for treatment responders (social phobia: Lundh & Öst, 2001; Mattia, Heimberg, & Hope, 1993) and may even predict how well treatment gains are maintained over time (e.g., e-Stroop effects predicted anxious thoughts both post-treatment and twenty months after initial testing following therapy for GAD; Mogg et al., 1995).
Thus, while more studies are needed evaluating a broader range of anxiety disorders and using a wider variety of paradigms (especially to address the question of whether change in measures tied to implicit social cognition mediates treatment response), the available evidence suggests a plausible functional relationship between change in automatic processing and symptom reduction.

**Linking Implicit Social Cognition Findings with Theoretical Models of Anxiety**

This updated review provides considerable support to prior theoretical proposals and reviews that emphasized the role of uncontrollable processing in anxiety disorders (e.g., Beck & Clark, 1997; McNally, 1995). At the same time, by examining results from a variety of new paradigms and a larger evidence base, we can have more confidence (than did some prior reviewers; e.g., McNally, 1995) that unconscious processing of threat stimuli also occurs preferentially for anxious persons. This conclusion is consistent with some more recent formulations (e.g., Öhman & Mineka, 2001), though results do vary across disorders and tasks (e.g., Phaf & Kan, 2007, reach a different conclusion when examining subliminal e-Stroop results only). Also, critical aspects of unconscious processing, such as anxious persons’ awareness of their biased responding, have yet to be examined (see Teachman, Joormann, & Gotlib, 2009). Analogously, while the evidence reviewed here is promising with regard to the role of automatic processes in enhancing vulnerability and being causally related to anxiety, as proposed by Beck and colleagues (e.g., Beck et al., 1985; Beck & Clark, 1997), the evidence is too limited at this stage to draw firm conclusions. Further, with currently available paradigms, we do not yet have the methodological elegance required to test Beck and Clark’s (1997) specific predictions about changes over time in the relative balance of automatic and strategic processing in anxious responding. These advances await the next new wave of measures to capture implicit social cognition.

**Mood Disorders**

Mood disorders are characterized by a disturbance of mood (depressed, elevated, or irritable) as the primary symptom. Major depressive episodes involve depressed mood or loss of interest or pleasure in normal activities lasting for at least two weeks, along with a range of cognitive (e.g., impaired concentration), motivational (low), behavioral (e.g., withdrawal), and somatic (e.g., changes in sleep and appetite) symptoms. Manic episodes are defined by abnormally elevated or irritable mood lasting for at least one week (hypomanic episodes are less severe), and often include racing thoughts, impulsive
behaviors (e.g., shopping sprees), and reduced need for sleep. These mood episodes are the "building blocks" of the mood disorders, which are divided into depressive and bipolar disorders.

Inspired by Beck's cognitive model of depression, which hypothesizes that depression is characterized by a schema that biases individuals toward selective processing of negative information (e.g., regarding personal failure and loss), researchers have investigated implicit cognition in depression since the late 1970s (see Beck, Rush, Shaw, & Emery, 1979). Overall, the evidence for implicit association, priming, and attention biases in depression is much more mixed than it is for anxiety disorders, with considerable variation across tasks. In their review, Hartlage, Alloy, Vasquez, and Dykman (1993) argued that depression interferes minimally with most features of automatic processing, with the exception of automatic activation of self-referent depressive content. However, in a more recent review, Teachman, Joormann, and Gotlib (2009) proposed that the current evidence suggests that depression is characterized by uncontrollable processing of negative information (see also Hertel, 2004).

Further, Beevers (2005) has proposed a dual process model, which distinguishes between the associative (implicit) mode of information processing (which is quick, effortless, and relies on pre-existing associations in memory) and the reflective (explicit) processing mode (which is slow, effortful, conscious, and rule-based). According to this perspective, cognitive vulnerability to depression occurs as a result of negatively biased self-referent associative processing. Depending on one's available cognitive resources and the degree to which associative processing violates expectations and triggers the need for more effortful processing, corrective reflective processing may or may not be able to overcome the negative associative bias. A feedback loop may form between negative associative processing, the absence of reflective processing, and dysphoric mood, resulting in clinical depression.

**Group Differences and Predictive Validity**

**Implicit Association Tasks**

Surprisingly little research has investigated group differences in implicit associations for individuals with mood disorders, though there are a few exceptions examining implicit self-esteem in depression. Some research on implicit associations has supported Beevers's (2005) prediction of negative associative processing among groups that are vulnerable to depression. For example, Gemar, Segal, Sagrati, and Kennedy (2001) found that formerly depressed participants showed significantly
stronger implicit associations on the IAT between the concepts “me” and “negative” after a negative mood induction, relative to control participants. Moreover, this negative bias was similar to that shown by currently depressed participants who had not undergone a mood induction, suggesting that negative state affect can activate underlying automatic critical self-evaluations in vulnerable individuals. In an interesting contrast, De Raedt, Schacht, Franck, and De Houwer (2006) found that depressed participants showed more positive implicit self-esteem than control participants on the EAST and found no group differences in implicit self-esteem on the IAT. The authors suggest that this may be due to normal early development of a positive implicit self-concept that remains latent despite the development of explicit symptoms of depression later in life. Alternatively, the discrepant results across studies may be tied to the presence of a mood manipulation (that enhanced the expression of the implicit associations) in one study and not in the other study. With regard to predictive validity, Haefelf et al. (2007) found that implicit but not explicit self-esteem predicted response to a lab stressor in unselected college students. Overall, few studies have been conducted examining implicit associations in depressed groups, and results are somewhat inconsistent. The presence of negative state affect or stressors may turn out to moderate the strength of negative implicit associations in depression, although clearly more work is needed.

**Priming Tasks**

Priming effects have been hypothesized to reflect difficulties with inhibition of task-irrelevant emotional material and have been linked with rumination, a key feature of depression (Joormann, 2006). In Beever’s (2005) dual process model, rumination is seen as a reflective processing strategy that may enhance the activation of task-irrelevant material and inadvertently amplify the effects of negative automatic processing (which may be detected through priming effects). Several studies have found that depressed individuals are more likely to show priming effects for negative stimuli than non-depressed individuals. For example, Bradley, Mogg, and Millar (1996) found that depressed participants showed larger subliminal and supraliminal priming effects following depression-relevant compared to neutral words on a lexical decision task, while control participants did not show this difference. Using the same task, these researchers also found that depressed participants showed greater subliminal and supraliminal priming from depression-relevant words than either a healthy control group or a group of individuals with anxiety disorders (Bradley et al., 1995). Similarly, Watkins, Vache, Verney, and Mathews
(1996) found that following conceptual encoding of valenced words, depressed participants produced more of the negative previously studied words in response to free association cues, while control participants produced more of the positive previously studied words.

Researchers have also found significant affective priming effects in depressed participants using more self-relevant paradigms. For example, Power, Cameron, and Dalgleish (1996) asked both depressed and non-depressed individuals to read a brief emotional prime and then evaluate as quickly as possible whether or not a target trait adjective was applicable to themselves. Depressed, but not control, participants showed faster responding and a higher rate of endorsement of negative adjectives when these adjectives were preceded by a negative emotional prime. Furthermore, Goeleven, De Raedt, Baert, and Koster (2006) used a negative priming paradigm, in which participants were asked to ignore a previously presented distracter picture of a face making a negative emotional expression, to examine group differences in affective priming. They found that depressed participants showed interference effects when making valence judgments of a target picture of a face presented later, but control participants did not. In a similar negative priming paradigm that used positive and negative adjectives as distracters and targets, Joormann (2004) found that non-dysphoric and never-depressed control participants showed negative affective priming (interference from a previously-presented, to-be-ignored negative prime on subsequent responding to a negative target). However, dysphoric and formerly depressed participants showed reduced negative priming effects; that is, responding to a negative target was quicker when it was preceded by a to-be-ignored negative prime (relative to the control groups).

Overall, these results suggest that depressed individuals may have more difficulty inhibiting negative information, even when it is irrelevant to the task at hand, compared to non-depressed individuals. Such findings support claims that reflective processing may be unable to correct automatic biases when negative material is highly activated or especially self-relevant (e.g., Beevers, 2005).

However, it should be noted that other studies do not find significant group differences in priming effects. For instance, both depressed and non-depressed participants showed priming effects for neutral but not emotional words in a study that used subliminal primes and a lexical decision task (Matthews & Southall, 1991). Furthermore, depressed and non-depressed groups have not shown differences in priming of positive and negative words in two studies of implicit memory using word completion tasks.
(Danion, Kauffmann-Muller, Grangé, & Zimmerman, 1995; Denny & Hunt, 1992). Similarly, Bisson and Sears (2007) found that individuals with depression did not show larger priming effects for negative targets, compared to other semantically related ones, with or without a negative mood induction.

Overall, group differences in priming effects seem more likely when stimuli are personally relevant or elaboration of stimuli is encouraged (e.g., through conceptual encoding of items). However, if extensive exposure and elaboration are necessary to observe effects, this raises questions about the extent the priming effects are really reflecting an automatic process (see Teachman et al., 2009). Also, when priming effects are observed, the contributing mechanisms are not yet clear. Priming effects may reflect an enduring processing style associated with depression, though other researchers (e.g., Watkins et al., 1996) have suggested that mood congruency (between the participant’s mood and valence of the stimuli) may also explain enhanced accessibility of negative material for depressed individuals.

Attention Bias Tasks

An attentional bias toward negative information has occasionally been found in depressed samples, using paradigms such as the e-Stroop task and the dot probe task (see review in Williams et al., 1996). Gotlib and Cane (1987) conducted an early study of e-Stroop effects in participants with clinical depression. They found that depressed participants showed increased interference from negative words, relative to control participants. Similarly, Segal, Hood, Shaw, and Higgins (1988) found an e-Stroop effect for primed self-descriptive words in depressed, relative to control, participants. However, these researchers did not find group differences for unprimed self-descriptive words. In another study using similar methodology but with only negatively valenced words as stimuli, they obtained the same results: primed words produced an e-Stroop effect in depressed individuals; unprimed words did not (Segal, Gemar, Truchon, Guirguis, & Horowitz, 1995). Notably, using masked stimuli presented at short exposure durations (14 ms), other studies have not found a bias for negative words in depressed participants (Bradley, Mogg, Millar, & White, 1995; Mogg et al., 1993). This suggests that elaborative processing of the negative content of self-relevant information is necessary for these stimuli to interfere with other cognitive tasks.

Studies using the dot probe task in depression initially produced mostly null findings. Although two early studies showed an attentional bias toward negative and social threat words in depressed,
compared to control, participants (Mathews, Ridgeway, & Williamson, 1996; Mogg et al., 1995), most other studies using unmasked stimuli produced no evidence of a bias for negative words in depression (e.g., MacLeod, Mathews, & Tata, 1986; Neshat-Doost, Moradi, Taghavi, Yule, & Dalgleish, 2000; Taghavi, Neshat-Doost, Moradi, Yule, & Dalgleish, 1999). Further, in the first study to examine pictures of threatening or sad faces in the dot probe task (which also included an eye tracking component), Mogg, Millar, and Bradley (2000) found no depression-specific biases in either reaction time or direction and latency of first eye movement. The use of masked stimuli presented at brief exposure times has also indicated no attentional bias at the subliminal level for depressed individuals (Mathews et al., 1996; Mogg et al., 1995).

Despite these null findings, more recent studies using longer stimuli exposure times and more ecologically valid stimuli, such as pictures of emotional faces, have found interesting differences in attention between depressed individuals and other groups. Gotlib and colleagues (Gotlib, Kasch, Traill, Joormann, Arnow, & Johnson, 2004) used a dot probe task with individuals who were diagnosed with clinical depression, social phobia, or no emotional disorder. They found that depressed participants were quicker to orient toward sad faces but slower to orient toward angry or happy faces, although they did not show between-groups differences from participants with social phobia within emotional face categories. Interestingly, severity of depressive symptoms was also related to a tendency to direct attention away from happy faces (see also Gotlib, Krasnoperova, Yue, & Joormann, 2004). Another study examining differences in attentional bias between currently and formerly depressed participants and never-disordered control participants found that both the currently and formerly depressed groups selectively attended toward sad faces, while the control group selectively attended toward happy faces (Joormann & Gotlib, 2007). These studies raise the possibility that depression is characterized more by a lack of attention to positive information, rather than excessive attention to negative information. To test this idea, Shane and Peterson (2007) had high and low scorers on the Beck Depression Inventory complete a modified dot probe task involving positively and negatively valenced pictures and words. They found that high scorers showed processing biases away from positive information (compared to neutral information) when stimuli were displayed for 200-500 ms. These participants also showed biases toward depression-
relevant stimuli. In contrast, low scorers showed a strong bias away from negative information while showing unbiased processing of positive information.

A few investigators have used visual search tasks to examine selective attention to negative information in depression. Studies using the face-in-the-crowd task, a variant of the visual search paradigm, have not found evidence that depressed individuals are distracted by negative facial expressions, relative to control stimuli (Karparova, Kersting, & Suslow, 2005). However, similar to the studies using the dot probe to investigate avoidance of positive stimuli, depressed participants were slower than control participants in responding to positive facial expressions, even though they did not differ in their speed of responding to negative facial expressions (Suslow et al., 2004). Adding to the mixed findings, Rinck and Becker (2005) used a visual search task involving depression-relevant words and found that depressed women were more distracted by these words than were control participants or participants with social phobia. Although the lack of studies makes conclusions tentative, visual search studies point to a likely attentional bias in depression; however, the extent the bias is specific to negative versus positive stimuli remains somewhat unclear.

While the vast majority of studies of cognitive biases in mood-disordered populations have examined depression, a few have investigated attentional bias relevant to bipolar disorder. French, Richards, and Scholfield (1996) administered an e-Stroop task to participants high in hypomanic symptoms. They found that these hypomanic participants showed e-Stroop interference for depression-related but not euphoria-related information, even when anxiety levels were taken into account. These researchers concluded that hypomanic personality traits may represent an attempt to cope with an underlying depressive cognitive vulnerability.

Overall, attention tasks generally show preferential attention to mood-congruent stimuli most consistently at long exposure intervals or when stimuli are self-relevant (see Mathews & MacLeod, 2005; Mogg & Bradley, 2005). Mixed results also seem tied to the critical question of whether depressed individuals are characterized more by avoidance of positive information or by excessive attention to negative information.

**Ties to Vulnerability and Causal Links**
Despite this large body of research investigating cognitive vulnerability factors involved in depression, few studies have used longitudinal or experimental designs that allow the examination of temporal prediction or causal relationships. In a notable exception, Haeffel et al. (2007) followed college students for five weeks after an initial IAT and self-report questionnaires to measure self-esteem and found that both implicit and explicit measures interacted with life stress to predict later depressive symptoms (although only the explicit measures were unique predictors). Similarly, Beevers and Carver (2003) measured attentional bias using a dot probe task with college students before and after a negative mood induction. They found that shifts in attention toward negative information following the mood induction interacted with life stress to predict increases in depressive symptoms seven weeks later. In addition, speed of mood recovery following the induction at Time 1 also interacted with life stress to predict depressive symptoms at Time 2. Extending this research to a diagnosed sample, Johnson, Joormann, and Gotlib (2007) administered an emotional face dot probe task to individuals experiencing a major depressive episode. Attentional biases to sad faces predicted change in symptoms of depression nine months later. These longitudinal studies suggest a temporal relationship between implicit association and attention biases and subsequent depressive symptoms, consistent with the possibility that these biases may enhance vulnerability.

Although studies of processing biases in individuals at risk for the development of a mood disorder (e.g., children of mood-disordered parents) are not a direct test of longitudinal prediction of clinical symptoms, they may suggest ties to vulnerability by showing biases in at-risk individuals. Along these lines, Joormann, Talbot, and Gotlib (2007) administered an emotional face dot probe task to daughters (ages 9-14) of depressed mothers. They found that these girls showed biased attention toward negative faces, compared to girls with non-depressed mothers. Similarly, Gotlib, Traill, Montoya, Joormann, and Chang (2005) administered an e-Stroop task to children of parents with bipolar disorder, following a negative mood induction. They found an attentional bias toward social-threat and manic-irritable words in these children, relative to children of never-disordered parents.

Finally, there is very limited published work directly addressing causal links between automatic processing biases and depression. However, promising related work by Joormann, Hertel, LeMoult, and Gotlib (2009) has shown that explicit instructions regarding word substitution strategies can enhance
depressed participants’ successful forgetting of negative words, and there is evidence that training attention biases away from threat content can influence subsequent negative mood in response to a stressor (e.g., MacLeod et al., 2002).

**Relation to Treatment**

Although we are not aware of any published research examining change in implicit associations following treatment for depression or other mood disorders, a few studies have looked at treatment effects on priming and attention biases. For instance, Dannlowski and colleagues (2006) examined inpatients with major depression, half of whom had a comorbid anxiety disorder, and compared them to healthy control participants on a subliminal affective priming task. Prior to treatment, they found strong interference effects from negatively valenced primes in the non-comorbid depressed group, but not in the healthy control group. However, after seven weeks of inpatient psychotherapy, recovered depressed individuals no longer showed interference from negative primes, compared to their previous performance.

Other research has shown that attention biases may also be attenuated following successful treatment for depression. For example, Gotlib and Cane (1987) found that group differences in e-Stroop interference from negative words in depressed patients versus non-depressed control participants disappeared after inpatient treatment. Similarly, Segal and Gemar (1997) administered an e-Stroop task before and after cognitive behavioral therapy for depression, accompanied by priming with self-descriptive or non-self-descriptive emotional phrases. After treatment, participants who were less depressed than they had been at pre-treatment showed less interference for negative targets primed by self-descriptive negative phrases (compared to targets primed by non-self-descriptive phrases), while participants whose depression levels did not change over the course of treatment showed the same amount of interference from negative self-descriptive primes as before treatment.

**Linking Implicit Social Cognition Findings with Theoretical Models of Depression**

In general, the available evidence for biased implicit processing in mood disorders is sparse and largely restricted to major depression. The few studies of implicit association biases that have been conducted have produced mixed results. However, priming studies have more often shown evidence of a bias, especially when material is self-relevant and when elaborative processing is encouraged. Similarly, attention bias studies are more likely to show group differences when self-relevant stimuli are presented.
at long exposure intervals, but the evidence for this type of bias is also somewhat inconclusive. Overall, however, the available evidence supports a likely role for some automatic processing of negative information in depression, particularly when opportunities for overriding biases with reflective processing are limited (e.g., when material is highly activated and/or self-relevant; in line with Beevers, 2005, and Hartlage et al., 1993), suggesting the processing may be somewhat uncontrollable under these circumstances (in line with Hertel, 2004, and Teachman et al., 2009).

**Eating Disorders**

Eating disorders (ED) are typified by an extreme disturbance in eating behaviors and attitudes. In the current chapter, we will focus primarily on evidence from tasks capturing implicit social cognition within the context of symptoms of anorexia nervosa (AN) and bulimia nervosa (BN), which are the two most well-established forms of EDs (Wilson & Pike, 2001). According to the DSM-IV-TR, AN is characterized by a refusal to preserve a healthy weight for one’s height and age, while at the same time experiencing extreme fears of either gaining weight or becoming fat. BN involves recurrent episodes of binge eating (i.e., eating very large volumes of food within a limited period of time), coupled with recurrent compensatory behaviors designed to prevent weight gain (e.g., self-induced vomiting, excessive exercise, fasting, etc.). In both AN and BN, self-evaluation is excessively tied to perceptions of one’s body size and shape. Note that a number of the studies in this section involve participants with symptoms of disordered eating, who did not qualify for an ED diagnosis (e.g., chronic dieters, etc.). Including individuals with subclinical symptoms is useful in this domain given evidence that ED symptoms may exist along a continuum (e.g., Lowe et al., 1996), and because there are high rates of individuals who do not meet full criteria for a specific ED (i.e., Eating Disorder Not Otherwise Specified; ED-NOS). For further information on implicit social cognition research associated with eating and weight among individuals without symptoms of disordered eating, see Wiers et al. (this volume).

According to the transdiagnostic model of disordered eating, a similar set of mechanisms are theorized to maintain both AN and BN (Fairburn, Cooper, & Shafran, 2003). In particular, individuals suffering from both disorders are thought to “overvalue” weight and shape concerns, which consequently become central to one’s self-concept (Fairburn, Cooper, Shafran, & Wilson, 2008). In addition to a preoccupation with eating and appearance, cognitive behavioral models emphasize the extent to which
individuals with symptoms of AN and BN attempt to control their weight through dietary restriction and/or compensatory behaviors (e.g., purging), which exacerbate bingeing (Fairburn et al., 2008). Eating disorders are therefore characterized by both an extreme desire to exert complete control over one’s shape and weight, as well as by the inability to realistically do so. Accordingly, although theoretical models are largely silent regarding the unique role that automatic cognitive processing biases play in disordered eating, it would be unsurprising if individuals with ED symptoms had difficulty with uncontrollable processing of shape, weight, and/or food cues given the centrality of control difficulties in eating disorders.

**Group Differences and Predictive Validity**

**Implicit Association Tasks**

Using an IAT, Vartanian et al. (2004, 2005) conducted a series of studies among restrained eaters (defined as chronic dieters) and unrestrained eaters (non-dieters) investigating implicit evaluations toward ED-relevant concepts, such as meal size, fattening foods, and body fat. Contrary to initial predictions, Vartanian and colleagues (2004) found that both restrained and unrestrained eaters had pronounced negative implicit evaluations toward fatness (relative to thinness). Vartanian et al. (2005) also found that, regardless of dieting status, participants were much quicker at associating large meals with words synonymous with “fat” (and small meals with words synonymous with “thin”), relative to the opposite pairing contingencies. Together, this research suggests that these particular associations do not differentiate between individuals high (versus low) in ED-relevant symptoms like chronic dieting.

At the same time, other studies indicate that certain types of implicit associations may be meaningfully related to symptoms of disordered eating, even among individuals who are not diagnosed with an ED. For example, Thomas, Judge, Brownell, and Vartanian (2006) found that two self-report measures of ED symptomatology were positively correlated with an IAT that assessed associations between anorexia and glamour (versus danger) among an unselected sample of undergraduate women. Furthermore, Ahern, Bennett, and Hetherington (2008) found that individuals who reported a greater “drive for thinness” on a self-report questionnaire were relatively faster to associate positive (versus negative) attributes with underweight models on an IAT. However, the sample used by Ahern et al. was unselected, and implicit association scores were unrelated to actual body dissatisfaction. Therefore,
interpretability of these findings for actual ED populations is somewhat limited (see also Ahern &
Hetherington, 2006).

Finally, Seibt, Häfner, and Deutsch (2007) recently investigated motivational reactions toward
pictures of food by using a version of the Affective Simon Task among individuals who either had or had
not eaten lunch. All participants were instructed to use a lever to pull a pictorial food stimulus toward
themselves as quickly as possible (approach), or push a pictorial food stimulus away from themselves as
quickly as possible (avoid). While both groups of students were faster at “approaching” (versus avoiding)
the food stimuli, this effect was especially pronounced among individuals who had not yet eaten lunch.
The results were comparable when researchers investigated this paradigm among individuals who
suffered from an ED (BN or AN), suggesting that the task was sensitive to hunger state for everyone,
regardless of ED status.

Together, it appears that implicit associations do not reliably differentiate between individuals high
and low in symptoms of disordered eating, but implicit associations may be meaningfully related to certain
ED symptoms (e.g., “drive for thinness”). More work must be conducted with diagnosed samples before
firm conclusions can be drawn, however.

**Priming Tasks**

Researchers have found that priming eating and/or weight-relevant concepts impacts subsequent
responding among individuals with symptoms of disordered eating (e.g., Ferraro, Wonderlich, & Johnson,
1997; McKee, Nhean, Hinson, & Mase, 2006). For example, Roefs, Stapert, Isabella, Wolters, and
Wojciechowski (2005) conducted an experiment to evaluate the hypothesis that individuals diagnosed
with AN or ED-NOS (compared to unrestrained eaters), would not display as strong a preference for
palatable versus unpalatable foods. Primes included fattening palatable foods (e.g., chocolate), non-
fattening palatable foods (e.g., strawberries), and non-fattening unpalatable foods (e.g., radishes).
Targets included generally positive words and generally negative words. As expected, within the AN
group, palatability of the food did not appear to impact responses toward the targets. Meanwhile, within
the healthy control group, primes did influence responses toward the targets insofar as this group tended
to be quicker on congruent trials (e.g., palatable-positive and unpalatable-negative), relative to
incongruent trials. This research suggests that there are differences in the ways in which food cues are
processed as a function of ED status, with palatability perhaps being a less important prime for individuals suffering from AN as it is for other people (Eertmans et al., 2001; Pliner & Mann, 2004).

Research also indicates that there is a relationship between symptoms of EDs and biased processing of ED-relevant stimuli on primed lexical decision tasks (e.g., Ferraro, Andrews, Stromberg, & Kristjanson, 2003). For example, Meijboom, Jansen, Kampman, & Schouten (1999) asked female participants high on symptoms of restrained eating to complete a subliminal lexical decision task to investigate the hypothesis that self-esteem is related to concerns with body weight and shape. Specifically, after low self-esteem was primed, Meijboom et al. (1999) found that there was a significant relationship between low self-esteem and “overconcern” tied to weight/shape. In other words, after the priming procedure, the accessibility of subliminally presented stimuli related to weight and shape were enhanced. On the supraliminal lexical decision task, this effect was not evident. Also, Cassin, von Ransom, and Whiteford (2008) found that individuals with subclinical ED symptoms (high thin-ideal internalizers) did not exhibit cognitive biases tied to disordered eating, even after being primed with pictures of attractive models. This suggests that ED symptom severity may need to be sufficiently high in order for priming paradigms to impact subsequent cognitive processing.

Attention Bias Tasks

Dobson and Dozois (2004) identified 28 empirical studies in a meta-analysis examining e-Stroop interference effects tied to food- (“Food e-Stroop”) and body-related words (“Body e-Stroop”), from approximately 1935 to the end of May 2001. They primarily limited their review to populations exhibiting symptoms of EDs, including dieting/restrained eaters, as well as individuals diagnosed with AN and BN. Their results indicated that individuals diagnosed with BN consistently demonstrated biases in attention on the e-Stroop across a variety of stimuli (e.g., Food e-Stroop, Body e-Stroop, and the Classic Stroop, which is not ED-relevant), but that attentional biases among individuals with AN were limited to weight/body-related stimuli. Furthermore, the authors concluded that there was not sufficient evidence to support the presence of an attentional bias among dieting/food-restricting participants (see also Lee & Shafran, 2004, and Lokken, Marx, & Ferraro, 2006, for the importance of symptom severity in the expression of e-Stroop interference effects). In contrast, results from Johansson, Ghaderi, and Anderson’s (2005) meta-analysis suggests that there may be small e-Stroop interference effects for body-
and food-related words among individuals who are not diagnosed with an ED but who are overly concerned with eating and body image. Thus, while attentional biases as assessed with the e-Stroop are clearly implicated in diagnosed ED populations, the extent to which these biases are present in subclinical populations awaits further testing.

In general, research also suggests that individuals suffering from disordered eating show biases on the dot probe detection task (though see Boon, Vogelzang, & Jansen, 2000), but there appear to be a number of critical moderators of the bias effects, including type and valence of the stimuli. For example, Shafran, Lee, Cooper, Palmer, and Fairburn (2007) conducted two studies to investigate attentional biases tied to so-called “positive”, “negative”, and neutral eating- and shape-relevant stimuli, and neutral weight stimuli (as compared to non-ED relevant stimuli). They compared women diagnosed with an ED to female healthy control groups (Studies 1 and 2), and to women with high levels of anxiety (Study 1). In the first study, researchers found that patients with EDs responded more rapidly toward a probe that appeared in the same spatial location as negative eating (e.g., high calorie food eaten in an uncontrolled manner) and neutral weight (e.g., scales or people being weighed) stimuli, and more slowly to a probe following positive eating (e.g., low calorie food eaten in a controlled manner) stimuli. Their second study largely replicated these results, except researchers found that individuals diagnosed with an ED also displayed an attentional bias toward shape stimuli that was negative (e.g., plumper bodies or body parts) or neutral (e.g., body parts not readily associated with weight/shape). Notably, these biases were greater for individuals diagnosed with EDs, relative to the anxious and control groups, although this was not true for the positive and neutral shape stimuli. Partially in line with these results, Rieger et al. (1998) found that participants diagnosed with an ED (in comparison to healthy control participants) were more likely to attend away from positive words signifying a thin physique, and there was a trend for them to attend toward negative words signifying a large physique. Consistent with Vitousek and Orimoto (1993), Rieger and colleagues (1998) suggested that patients with EDs preferentially process information that is congruent with underlying schemas (e.g., attending toward negative eating/shape stimuli), while being resistant to “counterschematic” information.

Placanica, Faunce, and Job (2002) also investigated attentional biases using a dot probe task among individuals who scored high (versus low) on symptoms of disordered eating. Individuals high in
ED symptoms were more likely to exhibit an attentional bias toward low-calorie food words, but only in a condition where they had not fasted. Authors interpreted these results to mean that as hunger increased, individuals high in ED symptoms shifted their attention away from low-calorie, and toward high-calorie foods. Finally, using a visual search task, Smeets, Roefs, van Furth, and Jansen (2008) recently found that individuals diagnosed with an ED, as compared to healthy control participants, displayed evidence of heightened distraction by high-calorie food cues on a visual search task, and enhanced orienting toward body-relevant cues (authors did not differentiate between the valence of body-relevant stimuli).

At this point, more work is needed to clarify when the valence, stimulus type, and directionality of attentional biases will be differentially related to ED symptoms. For example, on the dot probe task, Shafran et al. (2007) suggested that attentional biases toward shape stimuli are less robust than for eating or weight stimuli, and that shape stimuli may need to be personally-relevant in order for biases to manifest. Work by Placanica et al. (2002) also indicates that current hunger or dieting status is relevant for attentional biases toward food cues (see also Seibt et al., 2007), which may partially help to explain why some studies have found biases toward low-calorie stimuli (e.g., Placanica et al., 2002), while others have found attentional biases away from low-calorie stimuli (e.g., Shafran et al., 2007).

Notwithstanding, it appears that attentional biases associated with food- or body-relevant cues are meaningfully related to ED symptoms, particularly among diagnosed samples. Similar to anxiety disorders, there is also some evidence to suggest that attentional biases involve both initial orienting toward threat, as well as subsequent cognitive avoidance (Ben-Tovim & Walker, 1991; Lee & Shafran, 2008; McManus, Waller, & Chadwick, 1996; Rieger et al., 1998; Waller & Meyer, 1997; see also Lee & Shafran, 2004, for a review of information processing biases tied to ED symptoms).

Ties to Vulnerability and Causal Links

Currently, there is limited research on implicit social cognition outcomes as conferring vulnerability, or being causally related, to EDs.

Relation to Treatment

The evidence regarding treatment effects in EDs is mixed, and is mainly in the area of attention biases. Supporting the notion that attention biases are reduced following treatment, Cooper and Fairburn (1994) found that individuals with BN showed attenuated interference on an e-Stroop task that included
eating-, shape-, and weight-related words following successful treatment. Also, Ball, Mitchell, Touyz, and Griffiths (2004) investigated changes on a modified e-Stroop task among women diagnosed with AN. Results indicated that color-naming interference decreased following treatment for the food-related words, but not for shape-related words. Finally, Shafran, Lee, Cooper, Palmer, and Fairburn (2008) found that female participants diagnosed with an ED (versus healthy control participants) were more likely to display bias on a dot probe task toward eating, weight, and shape stimuli before treatment. Following treatment, these attentional biases were significantly weakened.

On the other hand, Carter, Bulik, McIntosh, and Joyce (2000) did not find treatment effects among women diagnosed with BN on an e-Stroop task as a function of word type (e.g., Body/Shape relative to Control words). In particular, treatment effects were not specific to the intended target of therapy (i.e., body/shape), although color naming latencies for all words were reduced as a function of treatment (see also Black, Wilson, Labouvie, & Heffernan, 1997). Similarly, Long, Hinton, and Gillespie (1994) found no significant change from pre- to post-treatment in color-naming latencies for body size or food-related words among a small group of individuals suffering from AN whose weight had normalized following a three-month treatment period. In fact, Lovell, Williams, and Hill (1997) found that when individuals had recovered from AN, they actually showed greater interference on shape-related stimuli, as compared to a small group of individuals who had recovered from BN. However, neither group differed significantly from one another or from healthy control participants on food-related words.

Thus, it is still an open question as to whether measures of attentional bias are sensitive to treatment effects among individuals with disordered eating, and whether these effects are specific to a given stimulus content, or whether they are more general.

**Linking Implicit Social Cognition Findings with Theoretical Models of Eating Disorders**

This research provides some support for cognitive behavioral models of eating disorders and suggests that under certain circumstances individuals with symptoms of AN and BN will preferentially process information tied to shape, weight, and/or food cues. However, clarifying the extent to which various features of automaticity are important for the expression of cognitive processing biases in AN and BN will require additional work. For example, the e-Stroop findings suggest that individuals with eating disorder symptoms may have difficulty with the unintentional and uncontrollable processing of ED-
relevant cues. However, research evaluating implicit associations, which also taps uncontrollable processing of threatening information, was much more mixed.

**Body Dysmorphic Disorder**

Body Dysmorphic Disorder (BDD) is classified as a somatoform disorder. Characterized by extreme distress and preoccupation regarding elements of one’s physical appearance (e.g., nose or skin), BDD symptoms cannot be better accounted for by concerns tied to one’s overall body weight or shape. Cognitive models suggest that among individuals with BDD, physical appearance is “overvalued” and plays a disproportionately large role in one’s self-concept (Buhlmann & Wilhelm, 2004). Researchers also stress that individuals with BDD may selectively attend to information relevant to their perceived defects, which in turn exacerbates the focus on even very minor flaws (Veale et al., 1996; Wilhelm & Neziroglu, 2002). While theoretical models have underscored the importance of cognitive factors for maintaining BDD symptoms, little emphasis has been placed on specifying which processing biases will occur relatively automatically. However, given the similarities between BDD and anxiety disorders (BDD is often conceptualized as an OCD spectrum disorder; Phillips, 1986), it seems plausible that uncontrollable and/or unconscious processing of threatening information is important for the expression, and perhaps development, of BDD.

**Group Differences and Predictive Validity**

**Implicit Association Tasks**

Researchers have recently demonstrated that implicit associations are meaningfully related to symptoms of BDD, though effects have not been consistent across BDD-relevant domains. For instance, Buhlmann, Teachman, Naumann, Fehlinger, and Rief (2009) used an IAT to evaluate implicit self-esteem, the importance of physical appearance, and an attractiveness stereotype (the association between competence and physical attractiveness). Participants included individuals diagnosed with BDD, individuals with subclinical BDD symptoms, and healthy control participants. As expected, individuals diagnosed with BDD (relative to control participants) displayed lower levels of implicit self-esteem. Meanwhile, the group with subclinical BDD symptoms had intermediary scores on the implicit self-esteem task (see also Buhlmann, Teachman, Gerbershagen, Kikul, & Rief, 2008, for similar implicit self-esteem findings). Individuals diagnosed with BDD (relative to the other two groups) also displayed stronger
implicit attractiveness stereotypes, such that they were more likely to associate attractiveness with competence. Furthermore, both the implicit self-esteem and implicit attractiveness stereotype task predicted distress and avoidance during an exposure to a mirror.

In contrast, Buhlmann et al. (2008, 2009) and Clerkin and Teachman (in press) found no significant differences between individuals with BDD symptoms (versus control participants) on an implicit measure of the importance (versus meaninglessness) of attractiveness. However, authors pointed to methodological constraints of the IAT to help explain these null findings; namely, it was difficult to find an appropriate relative comparison category for the importance of attractiveness (see discussion in Buhlmann et al., 2008, and Clerkin & Teachman, in press).

Priming Tasks

To our knowledge, there is no published work evaluating priming tasks among individuals with symptoms of BDD.

Attention Bias Tasks

Buhlmann, McNally, Wilhelm, and Irmela (2002) found that individuals diagnosed with BDD (relative to healthy control participants) displayed greater interference for both positive and negative words on an e-Stroop task, regardless of their relevance to BDD. Follow-up analyses indicated that these findings were most pronounced for positive words tied to BDD (e.g., “attractive”). Authors interpreted these results to mean that individuals with BDD may experience biases toward emotional cues generally, and that these effects are most pronounced for stimuli relevant to their specific concerns.

Ties to Vulnerability and Causal Links; Relation to Treatment

At this point, there is no published work (to our knowledge) tying implicit social cognition research to vulnerability or treatment effects within the context of BDD.

Linking Implicit Social Cognition Findings with Theoretical Models of BDD

This review provides support for cognitive models suggesting that individuals suffering from BDD overvalue the importance or meaningfulness of being physically attractive. Additionally, research investigating implicit associations and attentional biases suggests that individuals with BDD symptoms may be characterized by uncontrollable processing of threatening information, particularly when that information is tied to their specific areas of concern.
Schizophrenia and Other Psychotic Disorders

The primary psychotic disorders encompass a broad range of psychiatric illnesses, including schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder, and brief and shared psychotic disorders. In the present chapter, our focus is on implicit social cognition research primarily associated with schizophrenia, a disorder characterized by both “positive” and “negative” symptoms. Positive symptoms reflect mental phenomena that are in excess or that are a distortion of healthy functioning (e.g., delusions, which are fixed, false beliefs that are maintained in spite of indisputable evidence to the contrary, and hallucinations, which are distorted or false sensory perceptions). Negative symptoms are those that constitute a restriction of the intensity and/or range of healthy functioning, particularly in emotional processing (e.g., anhedonia, a difficulty experiencing pleasure).

It is important to note that much of the research regarding implicit processing within the context of psychosis has evaluated general cognitive or neuropsychological deficits (e.g., using the classic, non-emotional Stroop task as an index of executive processing; see Sitskoorn, Aleman, Ebisch, Appels, & Kahn, 2004; Szöke et al., 2008). For example, Fioravanti, Carlone, Vitale, Cinti, and Clare (2005) conducted a meta-analysis in which they found that individuals with schizophrenia tended to perform more poorly than healthy control participants on measures of executive functioning, such as the Stroop. There have also been a number of studies utilizing priming techniques, including lexical decision tasks, within the field of schizophrenia/psychosis research. However, most of this work has focused on investigating cognitive deficits, such as a generalized difficulty inhibiting task-irrelevant information in negative priming tasks (e.g., Moritz et al., 2001). In the current chapter our focus will be on research tied to content that is disorder-relevant (e.g., differential processing of positively or negatively valenced information), as opposed to processing difficulties more generally.

While biological models constitute the dominant theoretical perspective regarding the etiology of schizophrenia, researchers are beginning to emphasize psychological and social influences as important maintaining factors, particularly for positive symptoms of schizophrenia (Tarrier, 2008). Along these lines, cognitive behavioral models of schizophrenia highlight the importance of evaluating the relationship between one’s thoughts, feelings, and behaviors when responding to psychotic symptoms. Thus, one
goal of treatment might be to turn one’s attention away from auditory hallucinations or to re-evaluate maladaptive thoughts tied to delusional thinking (e.g., “the voices are completely out of my control”). Cognitive models of schizophrenia also emphasize the importance of enhancing one’s self-esteem or feelings of self worth (Hall & Tarrier, 2003). To our knowledge, there is little explicit discussion regarding the extent to which processing biases relevant to social cognition are occurring automatically versus strategically, although the ability to control one’s attention is clearly important in schizophrenia.

**Group Differences and Predictive Validity**

**Implicit Association Tasks**

Moritz, Werner, and von Collani (2006) used an IAT to evaluate implicit self-esteem, finding that patients with schizophrenia, particularly those experiencing current persecutory delusions, displayed lower levels of implicit self-esteem relative to healthy control participants. Interestingly, patients with schizophrenia who had persecutory delusions (versus those who did not) exhibited higher levels of explicit self-esteem. The authors suggest that persecutory delusions may partially function as a mechanism to counteract poor implicit self-esteem. However, with so little empirical work, it is difficult to draw firm conclusions regarding the role that implicit associations play in psychosis.

**Priming Tasks**

Because disorganized speech (comprised of tangential associations) is one of the common symptoms of schizophrenia, some researchers have hypothesized that the disorder is characterized by semantic “hyperpriming,” an exacerbation of the standard associations between words (see Maher, 1983; for reviews, see Goldberg & Weinberger, 2000; Minzenberg, Ober, & Vinogradov, 2002). However, in a meta-analysis of over 30 studies of semantic priming in schizophrenia, Pomarol-Clotet, Oh, Laws, & McKenna (2008) determined that direct semantic priming does not appear to be either increased or decreased in schizophrenia as a whole. Notwithstanding, studies of patients with schizophrenia who have symptoms of disorganized speech do tend to show evidence of increased priming of meaning, compared to healthy control participants and to schizophrenic patients without disorganized speech. Content of the word may be an important moderator. A study that looked at semantic priming of affectively valenced words (Rossell, 2004) found that neutral and happy word pairs showed significant
semantic priming, but fearful and sad word pairs did not; however, these results were the same for participants with schizophrenia as for healthy control participants.

Fewer studies have investigated affective priming in schizophrenia. In general, however, research suggests that individuals with schizophrenia may exhibit stronger affective priming effects toward certain types of stimuli, as compared with healthy control participants (see Hempel, Tulen, Van Beveren, & Hengeveld, 2005). For instance, Höschel and Irle (2001) utilized an affective priming task where participants were subliminally primed with faces conveying positive or negative expressions; then, following a mask, neutral facial expressions were displayed. Results indicated that the neutral faces were rated as significantly more unpleasant when following negative primes, as compared to positive and neutral primes, and that this effect was especially pronounced for patients with schizophrenia. The authors interpreted these findings to mean that the judgments of individuals with schizophrenia may be more strongly influenced by the automatic processing of negative emotional information. There is also some evidence that the judgments of individuals with schizophrenia, relative to healthy control participants, are less likely to be influenced by the automatic processing of positive facial expressions (Suslow, Roestel, & Arolt, 2003; see also Rossell, Shapleske, & David, 2001).

Current affective symptoms may be another moderator of priming effects in schizophrenia. Suslow, Droste, Roestel, and Arolt (2005) found that patients diagnosed with schizophrenia who did not currently have flat affect or anhedonia responded similarly to healthy control participants on an affective priming task. In general, both groups exhibited “reverse priming” or contrast effects whereby valence-incongruent face pairs facilitated processing and valence-congruent face pairs impeded processing of a target. However, schizophrenic patients with flat affect and with anhedonia showed a sensitivity to negative face primes, irrespective of the prime-target (in)congruency. Regardless of the valence of target faces, negative primes facilitated the evaluation of target faces within the patient group with flat affect, while negative primes interfered with the processing of target faces among patients with anhedonia (see also Suslow, Roestel, Droste, & Arolt, 2003). Notably, it is not entirely clear what factors moderate when congruency will facilitate versus interfere with processing target stimuli. Also, reasons for the discrepant results for persons with flat affect versus with anhedonia are not resolved, though may reflect differences in the salience of the emotional stimuli as a function of the different symptoms, given that anhedonia often
involves difficulty actually experiencing pleasure, while flat affect reflects a severe reduction in emotional expression (but the emotional experience may still be intact).

Attention Bias Tasks

Evidence regarding attentional biases for emotional information among individuals with psychotic symptoms has been somewhat mixed (see Garety & Freeman, 1999). This inconsistency may be partially due to the fact that schizophrenia and other psychotic disorders involve a very heterogeneous symptom profile, even when looking within positive and negative symptom clusters (Carpenter & Kirkpatrick, 2005). For example, Strauss, Allen, Duke, Ross, and Schwartz (2008) utilized two versions of an e-Stroop task among both "deficit" and "non-deficit" syndrome patients with schizophrenia, as well as healthy control participants. Patients with the deficit syndrome are typically conceptualized as being less able to experience all emotions, whereas individuals with schizophrenia without the deficit syndrome are less able to incorporate positive information. Results indicated that deficit syndrome patients (as compared to the other groups) failed to exhibit an attentional bias toward positive information; instead, these patients had difficulty disengaging from negative stimuli once their attention had been captured. Strauss and colleagues (2008) suggest that positive information may fail to capture the attention of deficit syndrome patients because it is inconsistent with their current mood state. These findings are also consistent with the priming studies (Suslow, Roestel, & Arolt, 2003; Suslow, Roestel, Droste, & Arolt, 2003; Suslow et al., 2005) described above, demonstrating that negative symptoms in schizophrenia may be associated with preferential processing of negative (but not positive) emotional cues.

In contrast, research by Muroi, Kasai, Uetsuki, and Suga (2007) suggests that individuals diagnosed with schizophrenia may not have an attentional bias for emotional information. Specifically, they found no significant group differences in response times or error rates across word conditions on an e-Stroop task, when comparing individuals diagnosed with schizophrenia to healthy control participants. Furthermore, McCulloch, Clare, Howard, and Peters (2006) evaluated responses on an e-Stroop task among patients with late-onset psychosis, older patients with depression, and healthy control participants who were matched for age. Again, results indicated no significant differences in attentional bias across the three groups.
Interpreting attentional bias results within the context of psychosis is further complicated by variable results depending on the presentation duration of threatening stimuli. For example, using a modified dot probe task, Arguedas, Green, Langdon, and Coltheart (2006) found evidence for a selective attentional bias toward threatening faces among a subclinical group of individuals prone toward delusions (versus healthy control participants who were not prone toward delusions). Interestingly, individuals high (versus low) in delusional-proneness exhibited an attentional bias toward threat-relevant (versus happy) faces at a brief 200 ms SOA only when the emotional faces were shown on the left side of the visual field. For longer SOAs (500 ms and 1250 ms), the attentional bias toward angry faces remained regardless of whether stimuli were presented in the left or right visual field. According to the authors, this suggests that selective attention toward threatening cues may be present among delusion-prone individuals at both automatic, as well as more strategic, stages of attentional processing.

Using the e-Stroop task, Bentall and Kaney (1989) also found that patients with persecutory delusions exhibited a selective attentional bias toward words signifying paranoid content (relative to control participants; see also Fear, Sharp, & Healy, 1996; Kinderman, 1994). This supports the notion that there may be an orienting or vigilance bias toward threat-relevant stimuli among paranoid individuals. However, there is also evidence from eye movement studies to suggest that within the context of psychotic symptoms, subsequent, more controlled visual attention may be selectively biased away from threatening cues (see Freeman, Garety, & Phillips, 2000; Green, Williams, & Davidson, 2003; Phillips & David, 1997), indicating a possible parallel with anxiety disorder models that also suggest attentional vigilance and then subsequent avoidance (e.g., Mogg et al., 2004).

In sum, the evidence for attentional biases tied to implicit social cognition among individuals with psychotic symptoms is mixed. The presence of affective symptoms (e.g., difficulty expressing emotions) and presentation duration of threatening stimuli appear to moderate the relationship between attentional biases and symptoms of psychosis, although it is difficult to draw firm conclusions given the relative dearth of studies and the conflicting findings to date.

**Ties to Vulnerability and Causal Links; Relation to Treatment**

To our knowledge there is no published implicit social cognition research evaluating vulnerability, causal links, or treatment effects among individuals with symptoms of psychosis.
Linking Implicit Social Cognition Findings with Theoretical Models of Schizophrenia

This review provides some support for cognitive behavioral models of schizophrenia, which contend that psychological and/or social factors, such as self-esteem, may contribute to the maintenance of schizophrenia (Tarrier, 2008). It is also clear that under certain circumstances, individuals with psychotic symptoms will preferentially process threatening social information at a relatively automatic level (e.g., Arguedas et al., 2006). However, drawing broad conclusions about the ways in which implicit social cognition theoretically informs cognitive models of schizophrenia is premature given that research is still very much in the early stages.

Personality Disorders

A personality disorder (PD) is diagnosed when a pattern of inner experience and/or behavior (i.e., cognition, affect, interpersonal functioning, or impulse control) deviates from the expectations of one’s culture, is inflexible and pervasive, causes significant impairment or distress, and begins by adolescence or early adulthood. PDs are divided into clusters defined by the involvement of odd/eccentric behavior (paranoid, schizoid, and schizotypal PDs), dramatic/emotional behavior (antisocial, borderline, histrionic, and narcissistic PDs), or anxious/fearful behavior (avoidant, dependent, and obsessive-compulsive PDs). The majority of implicit social cognition research in the area of PDs has been done on antisocial, borderline, and schizotypal personalities. Antisocial PD involves a pattern of disregard or violation of others’ rights, including failure to conform to social norms, unlawful behavior, lying, impulsivity, and lack of remorse. Borderline PD is a pattern of instability of interpersonal relationships, self-image, and affect. It is often characterized by impulsivity, efforts to avoid abandonment, chronic feelings of emptiness, and recurrent suicidal or self-mutilating behavior. Schizotypal PD is a pattern of social and interpersonal deficits, accompanied by cognitive or perceptual distortions (e.g., ideas of reference, paranoid ideation) and eccentricities of behavior (e.g., preoccupations with paranormal phenomena inconsistent with the norms of one’s subculture).

Similar to models of anxiety and mood disorders, cognitive models of PDs suggest that maladaptive schemas color an affected individual’s information processing style, resulting in selective processing of schema-relevant cues (see Beck, Freeman, & Davis, 2004; Linehan, 1993; Young & Lindemann, 1992). Examples of such maladaptive schemas include Entitlement/Grandiosity, the belief
that one is entitled to special privileges and is not bound by normal social conventions (theorized to be active in antisocial PD) and Defectiveness/Shame, the perception that one is defective, unlovable, and likely to be rejected (likely prominent in borderline PD; Young, Rygh, Weinberger, & Beck, 2008).

Schemas relevant to PDs would be expected to manifest as maladaptive implicit associations, enhanced emotional priming effects for disorder-relevant material, and attention biases for disorder-relevant cues.

**Group Differences and Predictive Validity**

**Implicit Association Tasks**

Two studies have examined implicit associations in violent offenders with antisocial personality characteristics. Gray, MacCulloch, Smith, Morris, and Snowden (2003) administered the IAT to a sample of individuals who had committed murder and had been admitted to a secure therapeutic community for male offenders with PDs. Half of their sample was diagnosed with psychopathy (which shares many features with antisocial PD), and half of their sample was diagnosed with other PDs. They found that murderers with psychopathy showed weaker implicit associations between the concepts of “violent” and “unpleasant” (relative to “peaceful” and “pleasant”) than murderers without psychopathy. These researchers also used the same IAT task with murderers who were high or low in psychopathy and a control group of offenders who had committed crimes other than murder (Snowden, Gray, Smith, Morris, & MacCulloch, 2004). Compared to non-murderers and murderers who scored low in psychopathy, murderers who scored high in psychopathy showed weaker unpleasant associations with violence. Interestingly, no group differences were found using explicit measures of attitudes toward violence.

Other research using the IAT has examined implicit shame in women with borderline PD. Rüschi, Lieb, and colleagues (2007) compared a group of women with borderline PD to control groups of women with social phobia or no emotional disorder on measures of explicit and implicit shame. Shame was hypothesized to be strongly related to borderline PD because of the connection between shame and suicidal intent, anger and impulsivity; all prominent features associated with borderline PD (see review in Rüschi, Lieb et al., 2007). Women with borderline PD endorsed more symptoms of shame and guilt on explicit measures than women in either control group. Mirroring this finding, they also showed stronger implicit associations of the self with shame-related (versus anxiety-related) words, compared to both control groups. In a separate study, however, these researchers found that women with borderline PD
and comorbid PTSD showed a more anxiety-prone (versus shame-prone) self-concept on the IAT, in contrast to women with borderline PD alone (Rüsch, Corrigan, et al., 2007).

Finally, using an unselected student sample, Campbell, Bosson, Goheen, Lakey, and Kernis (2007) found that narcissism symptoms showed a complex relationship with different aspects of implicit self-esteem, depending on whether agentic versus communal views of the self were being measured at the implicit level. Thus, at this stage, it seems clear that implicit associations can help differentiate among different aspects of personality pathology, supporting schema-based models, but variable results depending on the control group and specific IAT comparison categories make any general conclusions premature.

**Priming Tasks**

Studies of affective and semantic priming in people with antisocial PD have produced divergent results. Blair, Richell, Mitchell, Leonard, Morton, and Blair (2006) found that although individuals high in psychopathy did not differ in semantic priming effects from those low in psychopathy, they did show weaker affective priming for both positive and negative primes (see also Brinkley, Schmitt, & Newman, 2005). Blair and colleagues hypothesize that this reflects reduced sensitivity to affect in individuals with antisocial characteristics, even though processing of meaning is intact. Notably, these studies have not investigated priming with disorder-specific stimuli.

Schizotypal personality traits have also been associated with impaired emotion processing, as measured with an affective priming task. Kerns (2005) primed college students who scored high or low on measures of magical ideation and perceptual aberration (symptoms associated with schizotypy) with positive, negative, or neutral words prior to a lexical decision task. Individuals high in schizotypal symptoms did not show affective priming effects in reaction time or error rate, even though those low in schizotypal symptoms did.

The inflexible nature of PDs was examined in a study conducted by Bowles and Meyer (2008) with a sample of undergraduates who were assessed for avoidant personality features. Avoidant PD is a pattern of social inhibition, feelings of inadequacy, and extreme sensitivity to negative evaluation. Bowles and Meyer assigned participants to conditions in which they were primed by pictures reflecting secure or insecure interpersonal attachment or received no priming. When asked subsequently to evaluate an
ambiguous vignette describing a social interaction, participants high in avoidant traits made negative evaluations, regardless of priming condition, whereas those low in avoidant traits made negative evaluations only if primed with pictures depicting insecure attachment, suggesting more sensitivity to the prime condition. Again, the lack of priming effects for the high avoidant group suggests that some PDs are characterized by reduced sensitivity to affective or interpersonal cues.

One counterexample to this pattern comes from Domes and colleagues (Domes et al., 2006), who suggested that borderline PD may be marked by greater sensitivity to emotional primes. They investigated the role of inhibitory dysfunction in (unmedicated) women with borderline PD, compared to age-matched women without an emotional disorder. They found that women in the borderline group showed greater interference from negative primes than women in the control group did.

Attention Bias Tasks

Attention bias has been examined in borderline PD with mixed results. In the Domes et al. study (2006; described above), despite greater interference from negative words in the affective priming task, the women with borderline PD did not show any bias on an e-Stroop task that used negative and neutral words. In contrast, Sieswerda, Arntz, Mertens, and Vertommen (2007) found that a group of participants with borderline PD showed hypervigilance for emotional cues, as seen by interference on supraliminal and subliminal e-Stroop tasks. Individuals with another PD, an alternate psychological disorder, or no history of disorder did not show this bias, suggesting results were specific to borderline PD. Furthermore, the participants with borderline PD showed the strongest interference from disorder-relevant negative cues (e.g., “powerless” or “unacceptable”). Similarly, Sieswerda, Arntz, and Kindt (2007) found that individuals with borderline PD showed attentional biases on an e-Stroop task for both disorder-relevant and unrelated emotional stimuli, relative to control participants.

In addition, individuals with schizotypal personality traits may show an attention bias toward negative emotional stimuli. Mohanty, Heller, Koven, Fisher, Herrington, and Miller (2008) administered measures of schizotypy, anxiety, and depression to an unselected sample who then performed an e-Stroop task. Certain symptoms of schizotypy, such as perceptual disturbances or odd beliefs, were associated with attentional disturbance specific to negative words. Anxiety and depression symptoms did not mediate this relationship.
Interestingly, obsessive-compulsive PD (a pattern of perfectionism and preoccupation with order and control) and histrionic PD (a pattern of excessive emotionality and attention-seeking) have been hypothesized to show opposite patterns of attentional focus in the presence of potentially threatening stimuli. Individuals with obsessive-compulsive PD may be more likely to focus on small details, while individuals with histrionic PD may be more likely to focus on the "big picture" (see Maynard & Meyer, 1996). Yovel, Revelle, and Mineka (2005) used a visual processing paradigm in which participants were shown a large figure made up of smaller component figures and asked to respond to targets presented at either the local level (the component figures) or the global level (the overall figure). They found that participants high in obsessive-compulsive personality symptoms were more distracted than control participants were by the small, to-be-ignored details when they attempted to respond to the global figure, suggesting interference effects at the local level.

Ties to Vulnerability and Causal Links

We are not aware of any published studies that have manipulated implicit processing biases or have used longitudinal designs to examine effects on individuals with PDs. (However, see Locascio & Snyder, 1975, for an interesting early attempt to train selective attention to threat in order to influence symptoms of paranoia in an unselected sample.)

Relation to Treatment

Minimal work has examined change in attention biases following treatment for a PD, and we are aware of no published treatment research using implicit association or priming paradigms. Following an intensive three-year treatment program for borderline PD, Sieswerda et al. (2007) found that attentional biases on an e-Stroop task for disorder-relevant and emotional stimuli were no longer present for participants who had been judged as recovered from their disorder. However, they were still found in participants who had not recovered. It should be noted that the sample size in this study was small and showed significant attrition.

Linking Implicit Social Cognition Findings with Theoretical Models of Personality Disorders

Far more research in the area of PDs is needed, but the available evidence suggests some tentative conclusions for the three most studied disorders. Antisocial PD may be marked by implicit association biases and limited sensitivity to positive and negative affective primes. Borderline PD seems
to be characterized by implicit association biases and greater sensitivity to negative affective primes, and perhaps by attentional biases toward emotional cues as well. Schizotypal PD may be marked by limited affective priming effects and attentional biases toward negative information. Further, one study suggests that implicit PD-relevant biases may be changed following successful treatment. These findings are consistent with schema theories that propose selective processing of disorder-relevant cues, but do not yet permit fine-grained conclusions about the features of automaticity that are biased in PDs or the role of automatic biases in relation to vulnerability and response to treatment.

**Conclusion and Future Directions**

This review of the literature has simultaneously illustrated the tremendous progress that has been made in applying implicit social cognition principles and methods to psychopathology research and the considerable gaps in the literature. We are excited by the mounting evidence across paradigms and disorders that attest to the role that automatic biases play in the ways that individuals with psychological disorders view themselves and the social world. The findings appear most robust when examining group differences between clinical and non-clinical samples and prediction of symptoms and related behaviors, especially in the anxiety disorders. The evidence seems least certain when evaluating changes over the course of treatment, ties to psychopathology vulnerability, and causal links between biased processing and changes in symptoms. For these latter issues, despite some null findings, the biggest challenge is the paucity of studies that directly test these questions. Moreover, there is much we need to learn about variation across disorders, and about the effects of minor differences in paradigms. For instance, it is obvious that timing of stimuli presentation is a critical variable, both in terms of stimulus onset and duration of presentation, but it is difficult at this stage to provide general recommendations given the disparate results across studies, disorders, paradigms, etc.

Some of the challenges facing clinical researchers working in this field are likely the same as those facing all researchers trying to study implicit social cognition, regardless of the population or content focus. At the same time, we see some issues as being either unique to psychopathology research or perhaps just particularly tricky to manage. We conclude by highlighting just a few of these special issues and unresolved questions in the hopes that they can motivate future research in this area.
• The low reliability of many of the reaction time measures and frequent small effect sizes presents a considerable challenge to clinical researchers given the common focus on individual differences, and need for test-retest reliability to examine changes following treatment or vulnerability to disorder onset (e.g., Egloff, Schwerdtfeger, & Schmukle, 2005). Practice effects on the measures are a related concern, which has not received sufficient attention in the clinical literature. One reason these challenges are particularly salient for clinical researchers is the extreme difficulty of recruiting sufficiently large clinical samples to have adequate power to detect small effects and help mitigate the impact of low reliability.

• Another difficulty in applying some of the popular implicit social cognition measures to address clinical questions is that the measures are sometimes relative in nature (e.g., the IAT), so the target category of interest needs to be compared to an alternate target category. Interpretation of results from the measure is thus constrained by this comparison. For instance, while we spoke earlier in the chapter of “implicit spider fear associations” to simplify reporting of results, the many IATs examining spider fear associations have used variable comparison categories, so results across studies are not directly comparable. Moreover, selecting a reasonable comparison category is difficult for some of the complex associations posited by various psychopathology researchers. Take, for instance, the cognitive model of panic disorder (Clark, 1986), which proposes that panic occurs because benign bodily sensations are interpreted in a catastrophic way (e.g., a racing heart is misinterpreted as a sign of a heart attack). Attempting to measure catastrophic associations about bodily sensations is no easy feat, and our lab has had multiple failed attempts trying to show group differences on implicit association measures of this construct (e.g., Teachman et al., 2007). We think this is likely because there is no obvious comparison category for bodily changes. As newer implicit association measures are developed that do not require a direct target comparison category and that also have good psychometric properties (e.g., Affect Misattribution Procedure: Payne, Cheng, Govorun, & Stewart, 2005; Brief IAT: Sriram & Greenwald, in press) we look forward to increased flexibility with these tasks to assess some of the more nuanced clinical constructs proposed to operate at an automatic level.

• In addition to the issue of a good comparison or control category within a given task design, many of our studies are limited by the absence of other good controls. For instance, many of the studies discussed
in this chapter did not use an alternate psychopathology control group (helpful for showing that the bias measure is specific to a given clinical problem, versus common to psychopathology in general or to shared personality markers, such as neuroticism). Establishing specificity of the biases by showing discriminant validity has also been rare (e.g., showing that group differences exist on one disorder-specific task, but not on a comparably valenced, disorder-nonspecific task). Similarly, there has typically been insufficient attention paid to the role of comorbidity in our samples, yet we know that performance on many tasks is heavily influenced by this factor (e.g., Bradley, Mogg et al., 1995).

- Just as a given person often does not exhibit symptoms of only one disorder, the implicit social cognition measures themselves are not process-pure. The measures capture different aspects of automaticity (e.g., unconscious, uncontrollable, etc.) with variable degrees of effectiveness and the summary score for a task frequently does not disentangle the various processes that may influence task performance (e.g., role of guessing versus ability to overcome an automatic bias; see Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005, and Deutsch & Gawronski, 2009). This issue is especially important in the context of treatment, where clinicians hope that therapy can influence more than just the tendency to guess a certain way, but can actually help clients reduce activation of a bias or learn how to overcome maladaptive automatic responses when they do occur.

- Little is understood about the influence of state affect on some of the popular implicit social cognition measures (see Schmukle & Egloff, 2004, who found no effect of a state anxiety manipulation on an IAT in an unselected sample). When working with a population that has a clinical disorder (relative to a non-clinical comparison group), researchers need to contend with group differences tied to trait negative affect of course, but often differences in state negative affect are important as well. These may be present naturally (e.g., due to fluctuating symptoms), or groups may show differential responsiveness to a mood manipulation or stressor. Based on past research in the attention and memory bias literatures, it is not yet clear under what conditions state negative affect will suppress (e.g., Mathews & Sebastian, 1993) versus enhance (e.g., Chen, Lewin, & Craske, 1996) bias effects. Learning how these state/trait interactions influence various implicit social cognitive processes will be essential to understanding the real-world implications of these findings for emotion regulation, self-evaluation and social interactions.
We started this chapter commenting on our enthusiasm about the signs of increased dialogue between clinical and social cognition researchers. Using methods and theories across sub-disciplines is likely to benefit everyone. Along these lines, we see a number of possible lessons we can learn from the broader social cognition field. For instance, there has been a lot of recent interest in trying to determine potential moderators of relationships between implicit and explicit measures (e.g., Nosek, 2005), yet we know little about whether these will operate in similar ways with clinical populations. Social desirability is just one example of a commonly cited moderator in social cognition research. Would we expect it to operate the same way in a sample with schizophrenia that has deficits in executive functioning or distorted views of reality? What about in a sample with major depression or social anxiety, who often show reversals of the self-serving biases typical of psychologically healthy populations (e.g., Arkin, Appelman, & Burger, 1980)? Is there an equivalent clinical moderator analogous to motivation to control prejudice (see Plant & Devine, 1998)? Comparable questions surround the mechanisms to alter implicit associations as compared to explicit propositions (see Gawronski, & Bodenhausen, 2007). Much exciting work has occurred in recent years in the social cognition field to address these questions of moderators and mediators; how these findings can inform models of psychopathology is an important but unresolved question. On the flip side, clinical treatment studies and evaluations of predictive validity of symptoms can provide some informative lessons to social cognition researchers who want to learn more about the prediction of complex behaviors outside the laboratory and the durability of change in implicit social cognitive processes.

We look forward to watching this research area grow in the coming years, as clinical and social cognition researchers work together to start addressing these special challenges and unresolved questions. This is a stimulating time for researchers working across these two fields as new paradigms are coming on line that can help address some of the limitations of prior methodologies, and our ability to work across geographic and linguistic divides has made it easier to collaborate, share methods and ideas, and confront the challenges of recruiting clinical samples.
Clinical Applications: Chapter References


Change in implicit fear associations following treatment. *Journal of Abnormal Psychology, 112*, 100-
109.


disorder memoirs on readers' eating attitudes and behaviors. *International Journal of Eating
Disorders, 39*, 418-425.

task may indicate strategic but not automatic attentional bias. *Behaviour Research and Therapy, 35*,
131-144.

specific phobia. *Behaviour Research and Therapy, 35*, 29-34.

Vartanian, L. R., Herman, P., Polivy, J. (2005). Implicit and explicit attitudes toward fatness and thinness:
The role of the internalization of societal standards. *Body Image, 2*, 373-381.

Vartanian, L. R., Polivy, J., Herman, & C. P. (2004). Implicit cognitions and eating disorders: Their

disorder: a cognitive behavioural model and pilot randomised controlled trial. *Behavior Research and
Therapy, 34*, 717-729.

Vitousek, K., & Orimoto, L. (1993). Cognitive-behavioral models of anorexia nervosa, bulimia nervosa,
and obesity. In K. S. Dobson, & P. C. Kendall (Eds.), *Psychopathology and cognition* (pp. 191–243).

Vythilingam, M., Blair, S., McCaffrey, D., Scaramozza, M., Jones, M., Nakic, M., Mondillo, K., Hadd, K.,
attention in post-traumatic stress disorder: A help as well as a hindrance? *Psychological Medicine,
37*, 1445-1455.


Steketee (Eds.), *Cognitive Approaches to Obsessions and Compulsions—Theory, Assessment, and
Treatment* (pp.203 – 214). Oxford: Elsevier Science Ltd.


