The interplay and effectiveness of implicit and explicit avoidant defenses

Michael J. Marks¹ and Amanda M. Vicary²

Abstract
Individuals high on attachment avoidance are uncomfortable with thoughts of separation and loss. The goal of this research is to answer questions about the efficacy and interplay of the explicit (conscious) and implicit (preconscious) components of mental defenses designed to avoid uncomfortable thoughts. We manipulated the presence of subliminal attachment threat primes and participants’ awareness of those primes. While undergoing condition-specific threat manipulations, participants completed measures designed to measure attachment system activation. Avoidant participants who were aware of genuine attachment threat primes behaved defensively, whereas avoidant participants who were given false warnings of attachment threat primes did not. Results suggest that avoidant defenses operate on both implicit and explicit levels and are resilient to false activation.

Keywords
Attachment, attachment system activation, avoidance, avoidant defenses, defense mechanisms, psychology

Mental defenses are one of the fundamental concepts of Bowlby’s (1969/1982) attachment theory. According to attachment theory, the attachment behavioral system, an
evolved psychological mechanism, motivates humans (and some nonhuman species) to seek proximity to close others when faced with various threats. Although this system likely emerged to keep infants in close proximity to caregivers (Bowlby, 1969/1982), the attachment system persists into adulthood (Fraley, 2002) and manifests when adults are threatened as well (Hazan & Shaver, 1987). Understanding the nature of defenses and their role in attachment theory has been an important and long-standing research focus and is still often debated by contemporary attachment researchers. One such body of literature concerns a subroutine of these defenses concerning attachment-specific threats (threats related to separation and loss), and how attachment-related defenses operate in individuals high on attachment avoidance, who inhibit proximity seeking and attempt to handle distress autonomously.

Some evidence suggests that defensive strategies serve to protect avoidant individuals’ vulnerable, inner selves from various sources of threat, and if these defenses lapse, become fatigued, or fail, the attachment system can become reactivated, leading to rebounds of intrusive thoughts, heightened distress, and negative emotions (e.g., Ein-Dor, Doron, Solomon, Mikulincer, & Shaver, 2010; Mikulincer, Dolev, & Shaver, 2004; Wijngaards-de Meij et al., 2007). Other evidence suggests that the defenses of avoidant individuals render them robust against attachment system activation and can be effective in preventing attachment-related representations from being activated (e.g., Fraley & Brumbaugh, 2007; Fraley, Garner, & Shaver, 2000). This discrepancy may stem from the varied conditions under which threats are presented to and perceived by participants in past research. To date, little has been done to resolve these inconsistencies. The purpose of this research is to integrate these two sets of findings and, in the process, further advance our understanding of how the attachment system functions.

Resolving these inconsistencies is a means to answering questions about the efficacy of defenses and vulnerability to threats, including questions about what kinds of processes are involved in avoidant defenses (e.g., how effective are conscious and preconscious defenses?) and whether these processes are related in any way (e.g., are there “weaknesses” in one process that may be buttressed or exacerbated by the other?). To answer these questions, we assessed how variations in the presence and awareness of attachment threat primes affect the activation of attachment-specific defenses. These manipulations allowed us to determine whether attachment defenses operate most effectively at the explicit (conscious) level, the implicit (preconscious) level, or both. We begin with a brief overview of attachment mechanisms relevant to the present research. Next, we report the findings of an experiment designed to investigate the interplay between the implicit and explicit components of avoidant defenses. Finally, we discuss the implications of the present findings for attachment theory and research.

A brief overview of relevant attachment theory and research

According to attachment theory, people develop mental representations, or working models, that contain “rules” or “schemas” that guide interpretations of behaviors and interactions in close relationships (Bretherton & Munholland, 2008). These working models vary with respect to two dimensions, anxiety and avoidance (Fraley & Waller, 1998). Individuals high on attachment anxiety strongly desire closeness and intimacy,
but view themselves as inherently unworthy of love and thus fear abandonment and rejection (Mikulincer & Shaver, 2003). Individuals high on attachment avoidance prefer physical, psychological, and emotional distance from others (Fraley & Davis, 1997), view others as untrustworthy (Fraley, Davis, & Shaver, 1998), and value self-reliance and autonomy (Collins & Feeney, 2000; Mikulincer & Shaver, 2003).

According to Social Defense Theory (Ein-Dor, Mikulincer, Doron, & Shaver, 2010), there are unique schemas related to each of the major attachment orientations that affect how individuals react to and cope with threats. For avoidant individuals, these schemas involve handling threats autonomously with little regard for seeking affiliation with or warning others and distracting themselves from, circumventing, or ignoring threat-related information. Attachment-specific threats are handled similarly, and when faced with threats of separation or loss, individuals high on avoidance attempt to deactivate the attachment system in an effort to block or inhibit the negative emotional states associated with negative attachment experiences such as abandonment or rejection (Fraley et al., 1998). Additionally, attachment system activation threatens highly avoidant individuals’ sense of independence and autonomy, interfering with their desire for self-reliance. Importantly, avoidant individuals exhibit a dissociation between conscious and unconscious levels of responding (Shaver & Mikulincer, 2002). This suggests that avoidant defenses operate on two levels, a conscious, controlled level and a more preconscious, schematic, less effortful level. We describe both of these levels in more depth below.

**Implicit defenses**

We refer to defenses that operate at the preconscious level as implicit defenses. Like any other schema, these processes generally operate without conscious effort and use relatively few cognitive resources (see Ein-Dor, Mikulincer, & Shaver, 2011), but their operation is generally measureable by indirect measures (such as response latency or memory tasks). Contemporary models of the attachment system assume that attachment-related representations are based on a network of excitatory and inhibitory neural connections resulting from the repeated engagement in attachment style-specific strategies (Shaver & Mikulincer, 2007). These hypothetical connections are thought to facilitate or inhibit the spreading activation of attachment-related thoughts and feelings, resulting in variations in the likelihood of attachment system activation. Highly avoidant adults are thought to possess a network of sparse attachment-related experiences that inhibit spreading activation of attachment themes (Fraley, 2007). Accordingly, avoidant individuals tend to experience difficulty recalling life experiences (Mikulincer & Orbach, 1995) or information (Fraley et al., 2000) characterized by negative emotions or themes of loss. Even when motivated to recall such information, avoidant individuals’ ostensibly sparse network of attachment experiences results in difficulty in doing so (Fraley & Brumbaugh, 2007).

The sparsely connected network of negative attachment representations is one component of what Fraley and colleagues (Fraley & Brumbaugh, 2007; Fraley et al., 2000) refer to as preemptive strategies or defense mechanisms that serve to limit the amount of threat-related experiences that become encoded into memory. These mechanisms constrain the degree to which avoidant individuals can build and maintain a
rich network of detailed and coherent representations of attachment-related experiences, resulting in diminished accessibility, sensitivity, and attention to future closeness or emotion-related events. This process feeds back on itself, resulting in a maintained state of relative detachment.

The implicit defenses of avoidant individuals appear to be quite effective in the absence of overt or explicit attachment threats and when an ample amount of cognitive resources are available. Under such conditions, avoidant individuals have limited access to themes of separation and rejection, even when exposed to subliminal attachment threat primes (Mikulincer, Birnbaum, Woddis, & Nachmias, 2000; Mikulincer, Gillath, & Shaver, 2002). They also have greater access to positive self-representations and poorer access to negative self-representations relative to more secure individuals (Mikulincer, 1995).

Although generally resilient to unconscious activation of the attachment system, avoidant individuals are not necessarily immune to threats to attachment security, such as the loss of a loved one. A threat appraisal can lead to preconscious activation of the attachment system, increasing accessibility to attachment-related themes and representations (Shaver & Mikulincer, 2008). For avoidant individuals, preconscious activation of the attachment system can lead to conscious thoughts of negative attachment representations such as separations, rejections, or punishments (Mikulincer & Shaver, 2003), thoughts of which avoidant individuals find discomfiting.

**Explicit defenses**

Once faced with a clear or explicit attachment threat (e.g., relationship dissolution), avoidant individuals engage in more conscious, controlled means of deactivating the attachment system, such as suppressing attachment-related thoughts and memories, suppressing emotional displays, and distancing oneself from threat- and attachment-related cues (Shaver & Mikulincer, 2002). Through these means, avoidant adults can consciously regulate inward and outward responses to threat (see Fraley et al., 1998). For instance, avoidant adults who were told to suppress thoughts of abandonment experienced fewer intrusive thoughts of separation and showed lower skin conductance levels relative to those in control conditions (Fraley & Shaver 1997). In an examination of accessibility to attachment-related themes, Mikulincer, Dolev, and Shaver (2004) tested the robustness of explicit avoidant defenses. Participants suppressed or did not suppress thoughts of a romantic breakup, then performed a Stroop color-naming task (Stroop, 1935) task under either a high or low cognitive load. When not cognitively loaded, highly avoidant individuals displayed shorter color-naming latencies for words related to negative self-concepts, indicating low accessibility to negative self-representations. Avoidant individuals also displayed inflated self-appraisals in threatening situations (Mikulincer, 1998), ostensibly in an effort to validate their sense of autonomous self-reliance.

Finally, avoidant individuals can distance themselves from situations that threaten to activate the attachment system. When faced with relationship threats, avoidant individuals cope by diverting attention away from the threatening stimuli (see Banse & Imhoff, 2013). To document this effect in a real-world setting, Fraley and Shaver (1998)
unobtrusively recorded the behavior of couples in an airport terminal who were undergoing temporary separations. Highly avoidant individuals tended to avoid touching, expressing negative emotions such as sadness or crying, and eye gazing upon separation, demonstrating their propensity to avoid anxiety-provoking situations.

In sum, defenses can operate on an implicit level, without awareness or intention (e.g., Fraley et al., 2000; Mikulincer & Orbach, 1995) and independent of motivation (Fraley & Brumbaugh, 2007), as well as on a more explicit level (e.g., Fraley & Shaver, 1997), requiring intent and sufficient cognitive resources to be effective (Mikulincer et al., 2004). However, previous research on defenses has only focused on one of these processes at a time. To our knowledge, no researchers have examined the interplay between both processes. In order to achieve this goal, then, we experimentally manipulated both the presence of a threat prime and participants’ awareness of that prime, which allowed us to examine how the implicit and explicit components of defenses work alone and in conjunction with one another. These manipulations should address questions concerning whether the attachment system needs to be activated to defend against threats and how avoidant individuals react to threat cues if a threat fails to manifest.

**Overview and hypotheses**

In the present experiment, we systematically varied the presence and awareness of attachment threat primes to determine under what circumstances avoidant individuals do and do not experience attachment system activation. We employed a 2 (Threat Prime vs. Neutral Prime) × 2 (Warned of Threat Prime vs. Unaware of Prime) between-subjects design, such that participants were subliminally primed with either attachment threat words or neutral words and were either told a threat was present or were not. The efficacy of defenses was assessed in terms of both how strongly the attachment system was activated and the positivity of participants’ self-representations. Attachment system activation was assessed with a lexical decision task and a Stroop color-naming task, and self-representations were assessed through an adjective check list (ACL) task.

The lexical decision task required participants to indicate, as quickly as possible, whether a string of letters was a word (e.g., loving) or a nonword (e.g., lpving). Response times in a lexical decision task are based on stimulus compatibility processes, such that the speed of word identification is a function of the amount of relevant information currently activated—shorter response latencies indicate greater accessibility. Thus, attachment-related words should be more accessible to those experiencing attachment system activation than to those who are not, and hence reaction times to attachment words should vary with attachment system activation (i.e., comparatively fast reaction times suggest attachment system activation, comparatively slow reaction times indicate nonactivation or defensiveness). Lexical decision tasks have been validated as an effective means for exploring attachment-related representations (e.g., Mikulincer et al., 2000).

The Stroop task, based on processes of response compatibility, required participants to indicate the color of stimulus words as quickly as possible. An active mental representation increases attention to stimuli congruent with that representation, interfering with color naming of representation-relevant words. That is, a high color-naming latency
(i.e., a longer reaction time) is interpreted as an indicator of information accessibility. Thus, attachment system activation should lead to longer color-naming latencies. Attachment-related variations on the Stroop task have also been validated as an effective means for exploring attachment-related representations (e.g., Edelstein & Gillath, 2008; Mikulincer et al., 2004).

For the ACL, participants are shown a series of positive and negative adjectives and indicate whether or not they feel each adjective describes them. This task was used as a measure of participants’ self-assessments. Because avoidant individuals tend to inflate self-views in response to threatening situations (Mikulincer, 1998), we define defensive behavior in the ACL as endorsing comparatively more positive and fewer negative adjectives.

Based on the theoretical and empirical evidence discussed thus far, avoidant individuals should be robust against attachment system activation upon exposure to subliminal threat primes due to their implicit defenses. If this is the case, participants, as attachment avoidance increases, should exhibit (a) slower lexical decision reaction times, (b) faster Stroop reaction times, and (c) inflated self-perceptions in the ACL relative to others when exposed to subliminal threat primes. Avoidant defenses should also operate effectively at the explicit level (Fraley & Shaver, 1997, 1998). In other words, avoidant individuals should engage in defensive behavior if they are aware that they will be exposed to subliminal threat primes. If this is the case, we should, as attachment avoidance increases, observe the effects listed above in participants who are made cognizant of an impending threat prime. If avoidant defenses operate effectively at both implicit and explicit levels, then both of the above effects should occur. (They are not mutually exclusive.)

The above effects would manifest in two-way interactions between attachment avoidance and either of the manipulated variables (e.g., the effect of awareness might depend on the level of avoidance). Although these two-way interactions are important, a particularly critical interaction is the three-way interaction between avoidance, awareness, and threat. This interaction would indicate that attachment system activation depends on whether or not a threat prime is present, whether or not the person thinks a threat prime is present, and the degree of avoidance of that individual. Based on the discussion of the efficiency of avoidant defenses, we suggest it would not be prudent for avoidant individuals to engage their explicit defenses frivolously. As such, we predict that in the bogus threat condition, avoidant individuals will not exhibit evidence of attachment system activation. Further, highly avoidant individuals should only guard against explicit threats if a threat prime manifests. As such, we predict that only participants aware of a threat prime will, as attachment avoidance increases, exhibit (a) slower lexical decision reaction times, (b) faster Stroop reaction times, and (c) inflated self-perceptions in the ACL relative to others.

**Method**

**Participants**

The sample consisted of 211 undergraduates (52% female and 69% Caucasian) from a large Midwestern university who participated in exchange for partial class credit. All
were native English speakers. Average participant age was 19.86 years (SD = 1.80, range 18–27).

**Materials and procedure**

We informed participants that the study involved completing a personality questionnaire followed by various word tasks. Upon granting informed consent, participants completed the Experiences in Close Relationships Scale–Revised (ECR-R; Fraley, Waller, & Brennan, 2000), a 36-item attachment measure that assesses the dimensions of anxiety and avoidance. Cronbach’s \( \alpha \) for the present sample was .92 for the 18-item anxiety scale and .93 for the 18-item avoidance scale. The correlation between anxiety and avoidance was .41, consistent with previous attachment research using the ECR-R.

Participants then completed the ACL, lexical decision task, and the Stroop task in counterbalanced order. The subliminal primes and instructions given to participants varied according to condition. In the **threat aware** condition, participants received subliminal attachment threat primes (e.g., abandon and breakup) and were made aware that they would be subliminally primed with such words before each trial as a way of making thoughts of separation and loss highly accessible. In the **bogus threat** condition, participants received the same instructions as threat aware participants, but received neutral word primes instead of attachment threat primes. In the **threat naive** condition, participants received subliminal attachment threat primes, but were not informed of the prime. Participants in this condition received instructions that a flash of characters (actually the mask) would prepare the monitor for presentation of each stimulus. Finally, **control** participants received neutral word primes and the “character flash” instructions as in the threat naïve condition. All tasks were programmed using E-Prime software running on a Dell Dimension desktop computer. Stimuli were presented on a 15-inch, 60 Hz color cathode ray tube (CRT) monitor in white text (unless noted) in Arial, 28-point font against a black background. Responses were entered via serial response equipment designed for use with E-Prime. Following the debriefing, none of the participants accurately reported any of the prime words used in the research.

In the lexical decision task, participants identified whether a letter string was a word as quickly as possible. There were 100 randomly ordered trials consisting of 10 proximity words, 10 distance words, 10 neutral words, 10 positive nonattachment words, 10 negative nonattachment words, and 50 nonwords. The proximity and distance words were used to assess access to attachment themes, and the nonattachment-related words were included as to ensure the results were attachment specific (i.e., to ensure results were not simply based on word valence). Each trial began with a fixation point (a plus sign) centered on the screen. After a random delay of 3000 to 7000 ms, the 17 ms condition-specific subliminal prime was presented, followed by the presentation of a 500 ms mask (XXXXXXXX). The target letter string was then presented. Responses were entered via a serial response box—participants pressed a button labeled “yes” if the letter string was a word and a button labeled “no” if the letter string was not a word.

In the ACL, we presented participants with 15 positive adjectives (e.g., sincere, honest, etc.) and 15 negative adjectives (e.g., unkind, selfish, etc.) in random order and instructed them to indicate whether they felt each adjective was representative of
themselves. These adjectives originated from a larger pool of 30 positive and 30 negative words selected from Anderson’s (1968) list of trait adjectives that we tested in a pilot study. In order to avoid ceiling effects in the present study, we removed any adjective that was endorsed by more than 95% of the total pilot sample. Nine positive and zero negative adjectives were removed. The remaining adjectives were grouped by valence and used as a pool from which 15 positive and 15 negative adjectives were randomly drawn for use in the task. Each trial began with the presentation of a fixation point in the center of the screen. Following a random delay between 3,000 and 7,000 ms, the 17 ms condition-specific subliminal prime was presented, followed by a 500 ms mask. After the mask, the adjective was presented until the participant responded. Responses were entered via a serial response box—participants pressed a button labeled “yes” if they felt an adjective was self-descriptive and a button labeled “no” if it was not.

In the Stroop task, we presented participants with colored words and instructed them to verbally identify the color of each word as quickly as possible. There were 30 trials. Ten words were attachment-related words concerning distance (e.g., leaving, abandon, etc.), 10 were attachment-related words concerning proximity (e.g., hugging, loving, etc.), and 10 were neutral words (e.g., wires, carwash, etc.). Target words were presented in one of four randomly selected colors (red, blue, green, and yellow) on a black background, with the stipulation that no color was repeated twice in a row. Priming was executed in the same manner as in the lexical decision task. Participants spoke their responses into a microphone connected to a serial response box. Participants’ color-naming latencies were measured in milliseconds. The number of responses featuring ironic slips (e.g., the participant said or started to say the actual word “abandon” instead of saying the color “yellow”) was also recorded.

For the lexical decision and Stroop tasks, reaction times for each participant were averaged and examined for outliers. Reaction times greater than 2.5 SDs from each participant’s unique mean were dropped, per common psychology convention. Remaining reaction times for correct answers were averaged according to category (i.e., proximity, distance, positive, negative, neutral, and nonwords). Mean reaction times to proximity and distance words were used as dependent variables. Mean reaction times to neutral words served as a control, and mean reaction time to positive and negative words were controlled for in analyses involving proximity and distance words, respectively. Due to a very low occurrence of ironic slips in the Stroop task (three total slips across all conditions), analysis on this variable was abandoned. For the ACL, the number of positive and negative adjectives endorsed were summed by valence and served as the dependent variables.

Because of the continuous nature of attachment scores, hierarchical linear regression analyses were conducted on each dependent variable. Main effects and any control variables were entered in Step 1, two-way interactions in Step 2, and three-way interactions in Step 3. Anxiety and avoidance scores were centered in relation to their means, and categorical variables were effect coded such that, for the threat variable, “threat” was coded 0.5 and “neutral” was coded −0.5 and for the awareness variable, “aware of threat” was coded 0.5 and “unaware of threat” was coded −0.5. Significant interactions involving continuous variables were interpreted following the guidelines provided by Aiken and West (1991). Across all results below, variance inflation factors did not
exceed 5 and tolerance statistics were not below .70, suggesting multicollinearity was not a factor.

**Results**

**Lexical decision task**

**Distance words.** The full model predicted 57% of the variance in how quickly words related to separation and loss were identified, $F(12, 198) = 42.96, p < .001, R^2 = .57$ (see Table 1). In Step 1, there were no main effects of the manipulated variables or attachment scores. In Step 2, there was a two-way interaction between anxiety and threat ($\beta = .14, p = .009$), such that highly anxious people took longer to identify distance words when primed with a threat. In Step 3, there was a three-way interaction between avoidance, awareness, and threat ($\beta = .16, p = .002$). Simple slope analysis revealed three statistically significant slope differences: In the threat aware condition, participants displayed longer latencies as attachment avoidance increased ($B = 89.33, p = .002$). This slope was significantly different from the slopes of the other conditions, including threat naïve ($B = 3.19, p = .801; p_{\text{slope difference}} = .001$), bogus threat ($B = 20.10, p = .119; p_{\text{slope difference}} = .006$), and control ($B = 3.80, p = .519; p_{\text{slope difference}} = .002$). See Figure 1.

**Proximity words.** The full model predicted 48% of the variance in how quickly words related to proximity were identified, $F(12, 198) = 31.08, p < .001, R^2 = .48$ (see Table 2). In Step 1, there were no main effects of the manipulated variables or attachment scores.
scores, and in Step 2 there were no two-way interactions. In Step 3, there was a three-way interaction between avoidance, awareness, and threat ($\beta = .21$, $p < .001$). Simple slope analysis revealed three statistically significant slope differences: In the threat aware
condition, participants displayed longer latencies as attachment avoidance increased ($B = 79.56, p = .004$). This slope was significantly different from the slopes of the other conditions, including threat naive ($B = 1.81, p = .882; p_{\text{slope difference}} = .001$), bogus threat ($B = 8.50, p = .495; p_{\text{slope difference}} = .004$), and control ($B = 8.63, p = .121; p_{\text{slope difference}} = .008$). See Figure 2.

To summarize the lexical decision task results, participants high on attachment avoidance displayed diminished access to attachment themes relative to others when a threat prime was present and they were aware of it. When unaware of threat primes, participants high on attachment avoidance did not appear to experience differential access to attachment themes relative to others.

**Adjective check list**

**Negative self-representations.** The full model predicted 24% of the variance in how many negative adjectives were endorsed, $F(11, 199) = 5.75, p < .001, R^2 = .24$ (see Table 3). In Step 1, there was a main effect of anxiety ($\beta = .37, p < .001$), such that highly anxious people endorsed more negative adjectives. In Step 2, there were no two-way interactions. In Step 3, there was a three-way interaction between avoidance, awareness, and threat ($\beta = -.14, p = .047$). Simple slope analysis revealed three statistically significant slope differences: In the threat aware condition, participants endorsed fewer negative adjectives as attachment avoidance increased ($B = -2.82, p = .014$). This slope was significantly different from the slopes of the other conditions, including threat naive ($B = -0.39, p = .453; p_{\text{slope difference}} = .014$), bogus threat ($B = -0.50, p = .317; p_{\text{slope difference}} = .024$), and control ($B = 0.18, p = .429; p_{\text{slope difference}} = .010$). See Figure 3.
Table 3. Regression analysis of identification of endorsed negative traits in the Adjective Checklist as predicted by manipulated variables and attachment scores.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\Delta R^2$</th>
<th>$\beta$</th>
<th>$B$</th>
<th>$p$</th>
<th>95% CI $B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>.19</td>
<td>.000</td>
<td>6.09</td>
<td>.000</td>
<td>[5.68, 6.51]</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td>.06</td>
<td>.18</td>
<td>.429</td>
<td>.264</td>
<td>[−.078, 1.69]</td>
</tr>
<tr>
<td>Anxiety</td>
<td>.37</td>
<td>.23</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threat</td>
<td>−.15</td>
<td>−1.01</td>
<td>.018</td>
<td>.184</td>
<td>[−1.84, 0.18]</td>
</tr>
<tr>
<td>Awareness</td>
<td>.02</td>
<td>.10</td>
<td>.805</td>
<td></td>
<td>[−.72, 0.94]</td>
</tr>
<tr>
<td>Step 2</td>
<td>.04</td>
<td>.075</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoid $\times$ Threat</td>
<td>−.07</td>
<td>−.49</td>
<td>.286</td>
<td></td>
<td>[−1.4, 0.42]</td>
</tr>
<tr>
<td>Avoid $\times$ Aware</td>
<td>−.10</td>
<td>−.68</td>
<td>.142</td>
<td></td>
<td>[−1.59, 0.23]</td>
</tr>
<tr>
<td>Anxiety $\times$ Threat</td>
<td>−.10</td>
<td>−.65</td>
<td>.161</td>
<td></td>
<td>[−1.57, 0.26]</td>
</tr>
<tr>
<td>Anxiety $\times$ Aware</td>
<td>−.04</td>
<td>−.30</td>
<td>.528</td>
<td></td>
<td>[−1.23, 0.63]</td>
</tr>
<tr>
<td>Threat $\times$ Aware</td>
<td>.00</td>
<td>0.04</td>
<td>.959</td>
<td></td>
<td>[−1.62, 1.70]</td>
</tr>
<tr>
<td>Step 3</td>
<td>.02</td>
<td>.120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoid $\times$ Threat $\times$ aware</td>
<td>−.14</td>
<td>−1.83</td>
<td>.047</td>
<td></td>
<td>[−3.65, −0.02]</td>
</tr>
<tr>
<td>Anxiety $\times$ threat $\times$ Aware</td>
<td>.09</td>
<td>1.26</td>
<td>.183</td>
<td></td>
<td>[−0.6, 3.11]</td>
</tr>
<tr>
<td>Total $R^2$</td>
<td>.24</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. $N = 211$. CI = confidence interval, RT = reaction time.

Figure 3. Predicted mean number of negative traits endorsed in the Adjective Checklist task for high and low avoidance individuals under various combinations of awareness and presence of threat.

Positive self-representations. The full model predicted 26% of the variance in how many positive adjectives were endorsed, $F(11, 199) = 6.29, p < .001, R^2 = .26$ (see Table 4). In Step 1, there was a main effect of anxiety ($\beta = −.39, p < .001$), such that highly anxious people endorsed fewer positive adjectives. In Step 2, there were no two-way interactions.
In Step 3, there was a three-way interaction between avoidance, awareness, and threat ($\beta = .19$, $p = .005$). Simple slope analysis revealed three statistically significant slope differences: In the threat aware condition, participants endorsed more positive adjectives as attachment avoidance increased ($B = 2.87$, $p = .002$). This slope was significantly different from those of the threat naïve condition ($B = 0.58$, $p = .060$; $p_{\text{slope difference}} = .018$), the bogus threat condition ($B = 0.11$, $p = .796$) and the control condition ($B = 0.01$, $p = .964$; $p_{\text{slope difference}} = .008$). See Figure 4.

To summarize the ACL results, participants high on attachment avoidance displayed less negative and more positive self-representations relative to others when a threat prime was present and they were aware of it. When unaware of threat primes, participants high on attachment avoidance did not appear to experience differential self-representations relative to others.

**Stroop task**

**Distance words.** The full model predicted 67% of the variance in how quickly the colors of words related to separation and loss were named, $F(12, 198) = 33.39$, $p < .001$, $R^2 = .67$ (see Table 5). In Step 1, there were no main effects of the manipulated variables or attachment scores, and in Step 2, there were no two-way interactions. In Step 3, there was a three-way interaction between avoidance, awareness, and threat ($\beta = -.14$, $p = .003$). Simple slope analysis revealed three statistically significant slope differences: In the threat aware condition, participants displayed shorter color-naming latencies as attachment avoidance increased ($B = -0.68$, $p = .001$). This slope was significantly different from the slopes of the other conditions, including threat naïve ($B = -5.81$, $p = .492$;
The full model predicted 52% of the variance in how quickly the colors of words related to separation and loss were named, $F(12, 198) = 17.63, p < .001, R^2 = .52$.

Proximity words. The full model predicted 52% of the variance in how quickly the colors of words related to separation and loss were named, $F(12, 198) = 17.63, p < .001, R^2 = .52$. See Figure 5.
In Step 1, there were no main effects of the manipulated variables or attachment scores, and in Step 2, there were no two-way interactions. In Step 3, there was a three-way interaction between avoidance, awareness, and threat ($\beta = -.17, p = .002$). Simple slope analysis revealed three statistically significant slope differences: In the threat aware condition, participants displayed shorter color-naming latencies as attachment avoidance increased ($B = -56.22, p = .007$). This slope was significantly different from the slopes of the other conditions, including threat naı̈ve ($B = -3.69, p = .695; p_{\text{slope difference}} = .004$), bogus threat ($B = -3.05, p = .748; p_{\text{slope difference}} = .004$), and control ($B = -2.79, p = .047; p_{\text{slope difference}} = .008$). See Figure 6.

To summarize the Stroop task results, participants high on attachment avoidance displayed diminished access to attachment themes relative to others when a threat prime was present and they were aware of it. When unaware of threat primes, participants high on attachment avoidance did not appear to experience differential access to attachment themes relative to others.1

Discussion

The goal of the present research was to help resolve inconsistencies in the literature regarding the efficacy of avoidant defenses. Because this inconsistency may stem from research examining different parts of the defensive process, we varied the presence and the awareness of an attachment threat and measured attachment system activation, allowing for a more inclusive look at avoidant defenses.

The lexical decision and Stroop task results revealed that avoidant participants did not exhibit implicit access to attachment themes in response to attachment threats and that they did not experience attachment system activation if they were aware of a threat.
Further, avoidant participants did not display longer response latencies in the bogus threat condition, suggesting they did not engage defenses upon the mere insinuation of a threat—the threat had to be present. Similarly, the ACL results revealed that highly avoidant participants endorsed fewer negative and more positive adjectives when faced
with explicit threats than those of other attachment orientations and did not exhibit implicit vulnerabilities.

**Implications of the present results**

The present results have several implications for attachment theory and research. First, we suggest they can serve as a starting point at which to resolve the discrepancy between research suggesting that avoidant individuals are robust against attachment threats and research suggesting avoidant individuals are vulnerable to such threats. In the present study, avoidant participants were more adept than others at separating out genuine threat primes from false alarms, and consequently did not react defensively upon the mere suggestion that an attachment threat prime was present. These findings suggest that if avoidant individuals sense an impending threat (e.g., a relationship’s decline, a loved one falling ill), their implicit defenses may help inhibit any negative experiences associated with the actual occurrence of the threat from being strongly encoded into their memories. In any case, whether or not the threat occurs, there should be little to no expenditure of resources to activate explicit defenses.

On the other hand, if a threat suddenly and unexpectedly arises (e.g., the sudden death of a romantic partner in a car accident), avoidant individuals may then have to rely on more controlled mechanisms of defense, such as suppression. In sum, we suggest avoidant individuals have two lines of defense: they are sensitive to genuine attachment threats and can “maneuver” around them to prevent any negative experiences from becoming encoded into memory, and they can suppress any negative memories that do become encoded. However, as past research suggests, when unable to keep encoded memories suppressed (due to a cognitive load, prolonged stress, or some resource-consuming activity), they may then suffer rebounds of these memories, leading to difficulty adjusting to traumatic events (e.g., Fraley, Fazzari, Bonanno, & Dekel, 2006; Mikulincer et al., 2004).

If this is the case, the apparent disparity in previous research on attachment defenses may be the result of focusing on different components of defense. It may well be the case that avoidant individuals can prevent many attachment-related experiences from being encoded into their memories. However, research that focuses entirely on conscious suppression of memories likely involves experiences that have already been encoded, thus bypassing a crucial component of defenses. Avoidant individuals may be very adept at filtering out most unwanted or unpleasant experiences from their memories, but if a research paradigm activates only the experiences that have made it through the filter, then avoidant individuals’ ability to protect themselves from attachment system activation is likely underestimated. Likewise, research that focuses only on the encoding portion of attachment defenses may overestimate avoidant individuals’ ability to keep the attachment system deactivated. In sum, we suggest the attachment system does not necessarily need to be activated in order for highly avoidant individuals to be robust against intrusive thoughts.

Second, although avoidant individuals may experience some vulnerability to rebounding or intrusive attachment thoughts, it is unlikely this activation is tantamount to that experienced by anxious individuals. Anxious individuals’ hyperactivating strategies lead them to be more sensitive to threats than avoidant individuals, which could
result in better and more accurate threat detection (see Ein-Dor & Perry, 2014), but could also lead to more false positives (see Fraley, Niedenthal, Marks, Brumbaugh, & Vicary, 2006) depending on the stimuli. Avoidant individuals, on the other hand, appear better able to separate real threats from false alarms. This ability may help avoidant individuals defend against unwanted activation of the attachment system. On the other hand, just because a threat is detected does not imply that an avoidant individual will address it; indeed, avoidant individuals may simply avoid or ignore the threat, which may reduce any unwanted feelings in the short term, but may also have a negative impact in the long run should the issue not be addressed.

At a conceptual level, these findings support a control system model of attachment system dynamics in adulthood, which assumes the presence of excitatory neural circuits that result from the repeated use of hyperactivating or deactivating strategies (see Mikulincer & Shaver, 2003). Specifically, the subjective appraisal of threats, not necessarily their actual presence, is a crucial facet to this type of model. Mikulincer and Shaver (2003) explain that threat appraisals are not only dependent on the presence of a threat, but on the perceiver’s subjective perception of the threat (which can be biased by excitatory and inhibitory neural circuits). The present results suggest that avoidant individuals appear unlikely to subjectively interpret an ambiguous or vaguely threatening event as an actual threat.

Strengths and limitations
Manipulating both the presence and awareness of attachment threats allowed for testing both implicit and explicit components of defenses. The key condition was the “bogus threat,” where participants were falsely led to believe attachment threat primes were present. This allowed us to test whether avoidant individuals would act defensively upon the mere suggestion of a threat or if there would actually have to be a genuine threat present before avoidant individuals engaged their defenses. However, the sample was relatively homogenous in age and ethnicity. Younger individuals may not have had as much relationship experience as older individuals, or may still be developing secondary attachment strategies. However, by adolescence, the years of neuroplasticity have already passed, and any associative connections created by early attachment experiences should be long completed.

Conclusion
The present research suggests that avoidant defenses operate at both the implicit (relatively automatic and unconscious) and explicit (controlled and conscious) levels. Although the implicit components of defense appear useful in allowing avoidant individuals to overlook situational stimuli, deeper, emotional threats require a more effortful defense, requiring awareness of the threat itself. Hopefully, this connection will prove to be fruitful and lead to wider understanding of attachment, avoidance, and the nature and purpose of avoidant defenses.

Authors’ note
Any materials pertaining to the present research may be obtained via request from the first author.
Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note
1. Should the control words not be included in Step 1 of the Stroop task regressions, the three-way interaction remains statistically significant, but the overall model does not. The pattern of slopes and slope differences is identical in both analyses.

References
Ein-Dor, T., Mikulincer, M., Doron, G., & Shaver, P. R. (2010). The attachment paradox: How can so many of us (the insecure ones) have no adaptive advantages? *Perspectives on Psychological Science, 5*, 123–141. doi:10.1177/1745691610362349


