Reality Monitoring in Adults Reporting Repressed, Recovered, or Continuous Memories of Childhood Sexual Abuse

Richard J. McNally, Susan A. Clancy, Heidi M. Barrett, and Holly A. Parker
Harvard University

People who report either repressed or recovered memories of childhood sexual abuse (CSA) may have deficits in reality monitoring—the process whereby one discriminates memories of percepts from memories of images. Using signal detection methods, the authors found that adults reporting either repressed or recovered memories of CSA were less able to discriminate between words they had seen from words they had imagined seeing than were adults reporting either never having forgotten their CSA or adults reporting no history of CSA. Relative deficits in the ability to discriminate percepts from images (i.e., low $d'$) were apparent on only some tests. The groups did not differ in their criterion—response bias—for affirming having seen versus imagined stimuli.

The controversy concerning alleged repressed and recovered memories of childhood sexual abuse (CSA) has been among the most bitter in the history of psychology (McNally, 2003b). Endorsing concepts such as repression and traumatic amnesia, some psychologists affirm the authenticity of many of these memories (e.g., D. Brown, Schefflin, & Whitfield, 1999), whereas others are skeptical (e.g., Loftus & Ketcham, 1994).

Skeptics have proposed mechanisms that might lead people to develop false memories of CSA (e.g., Lindsay & Read, 1994). One possibility is poor source-monitoring ability—the skill enabling people to identify the origins of the contents of their minds (Johnson, Hashtroudi, & Lindsay, 1993). For example, source monitoring comes into play when people attempt to determine whether they heard something on TV or read it in the newspaper. More specifically, people prone to develop false memories of abuse may have a deficit in reality monitoring—a form of source monitoring required for distinguishing mental contents arising from perception from those arising from imagination (Johnson & Raye, 1981). Reality monitoring enables people to answer questions such as “Did this happen to me? Did I experience this? Or did I just imagine or dream it?” People with poor reality-monitoring skills may have difficulty distinguishing memories of previous experiences from memories of experiences that were merely imagined. For example, some therapists who suspect a patient suffers from repressed memories of CSA will ask the patient to visualize scenarios of possible abuse. A patient with reality-monitoring deficits may subsequently confuse a memory of such visualized scenarios with a memory based on a genuine event. Indeed, some people who believe they harbor repressed memories of abuse arrive at this conclusion after experiencing nightmares, visual images, or bodily sensations that they interpret as perceptual-sensory fragments of buried memories of early trauma. The hypothesis about reality-monitoring deficits implies that people who report repressed and recovered memories are especially prone to interpret certain experiences (e.g., images, dreams) as memories of actual events; it does not imply that they have these experiences more often than other people do (e.g., elevated nightmare frequency).

Experiments on cognitive functioning in people reporting repressed and recovered memories of CSA have only recently begun to appear (for reviews, see McNally, 2003a; McNally, Clancy, & Barrett, 2004). Although researchers have yet to investigate possible reality-monitoring deficits in these individuals, they have studied propensity to form false memories in the laboratory. In one experiment, women reporting recovered memories of CSA had a higher rate of false recognition of nonpresented critical words in a variant of the Deese (1959)/Roediger and McDermott (1995; DRM) paradigm than did women who had never forgotten their abuse (Clancy, Schacter, McNally, & Pitman, 2000). Participants who mistakenly affirmed having seen a word that captures the gist of words they had seen may have made a reality-monitoring error. That is, they may have generated an image of the nonpresented critical word while processing semantically related words and then mistakenly “remembered” having seen the critical word. People who report having been abducted by space aliens also exhibit heightened propensity for false memory formation in the DRM paradigm (Clancy, McNally, Schacter, Lenzenweger, & Pitman, 2002). On the other hand, recovered memory participants did not exhibit a heightened tendency to affirm having experienced childhood events that they had previously visualized on an imagination inflation task (Clancy, McNally, & Schacter, 1999).

The chief purpose of the present experiment was to test whether participants reporting either repressed or recovered memories of CSA exhibit reality-monitoring deficits relative to participants reporting either continuous memories of CSA or to a comparison group of participants reporting no history of CSA. To accomplish this aim, we recruited adults from the community who either (a) believed they harbored repressed memories of CSA, (b) reported having recovered previously forgotten memories of CSA, (c) re-
ported never having forgotten their CSA, or (d) reported never having experienced CSA. We adapted the methods of H. D. Brown, Kosslyn, Breiter, Baer, and Jenike (1994). Participants were presented with a series of words and asked to compare the relative heights of the first and fourth letter of each word. Half of the words were printed on cards, and the other half were heard on audiotape. For audiotaped words, participants had to generate a visual image of the word to perform the letter-height task. Following this first trial block of words, we asked participants to identify which words they had seen (vs. which ones they had imagined). For the second and final trial block of words, participants knew in advance that they would be asked to identify what words they had seen. Providing participants with this foreknowledge enabled us to test whether their ability to discriminate between percepts and images improved.

Following H. D. Brown et al. (1994), we calculated each participant’s hit rate and false alarm rate. The hit rate was the proportion of words that had appeared on cards that the participant correctly affirmed he or she had seen, whereas the false alarm rate was the proportion of words that had been presented on the audiotape that the participant had incorrectly affirmed having seen.

Using signal detection methods, we tested three hypotheses. If participants reporting either repressed or recovered memories of CSA are characterized by reality-monitoring deficits, then they should have lower sensitivity (d’) for distinguishing percepts from images than should participants either reporting continuous CSA memories or reporting no history of CSA. That is, the lower a participant’s d’, the more difficult it will be for a participant to discriminate the products of imagination (words they had imagined) from those of perception (words they had seen). Just as some people are better able than others to distinguish a signal from a background of noise, some people are better than others at telling whether a mental content is a product of perception rather than imagination.

In addition to experiencing difficulty discriminating the products of imagination from those of perception (i.e., low d’), there is another independent mechanism through which a person might endorse false memories: response bias. When in doubt, some people may be more likely than others to conclude that mental contents are products of perception (not imagination), then they should have a lower criterion (C) for claiming that a word had been seen relative to participants reporting either continuous CSA memories or no abuse history. The lower a person’s threshold for affirming having seen a word, the more likely the person is to produce correct responses for those items that were visually presented but the more likely to produce incorrect responses (false alarms) for those items that had been produced in imagery.

If individuals reporting either repressed or recovered memories of CSA experience difficulty exploiting foreknowledge about the subsequent memory test, then they should not exhibit improvement in their d’ scores from the first to the second trial block, relative to participants reporting either continuous CSA memories or no abuse history. On the other hand, if those reporting repressed or recovered memories of CSA can enhance their encoding of source information (perception vs. imagination), then they ought to show improvement from the first to the second block.

Finally, we examined possible predictors of d’: visual imagery ability, proneness to dissociate, and proneness to become absorbed in fantasy. Each of these variables should be inversely correlated with d’. That is, a trait associated with vivid imagery (imagery ability, absorption) or that reflects proneness for disruptions in consciousness (dissociation) ought to make it difficult for individuals to discriminate mental contents arising from perception from those arising from fantasy. Indeed, Rassin, Merckelbach, and Spaan (2001) found that individuals reporting difficulty distinguishing dreams from reality had elevated scores on the Dissociative Experiences Scale (DES; Bernstein & Putnam, 1986). People who believe they harbor repressed memories of CSA likewise score higher on the DES than do people who report always having remembered their CSA (McNally, Clancy, Schacter, & Pitman, 2000).

In summary, we sought to determine whether adults reporting repressed or recovered memories of CSA experience difficulty distinguishing percepts from images and, if so, to characterize this difficulty by determining whether it arises from low sensitivity, response bias, or both. Note that there are two independent ways in which the “false memory perspective” could be confirmed: Repressed and recovered memory participants might either experience difficulty distinguishing the products of imagination from those of perception (i.e., low sensitivity) or they might have a response bias for affirming that mental contents result from perception (i.e., low criterion for affirming having perceived something rather than having imagined it). Evidence against the “false memory perspective” would entail these groups being indistinguishable from the comparison and continuous groups on both d’ and C.

Method

Participants

We recruited participants from the community by posting announcements in public places, including newspapers, that read:

- Were you sexually abused as a child?
- Do you think you might have been sexually abused?
- Do you have no history of childhood sexual abuse?

The notices stated that individuals responding “yes” to any of these questions “may be eligible to participate in a research study at Harvard University concerning memory and trauma.”

Individuals responding to our notices received a phone screen to determine their eligibility. Upon arriving at the laboratory, they provided written informed consent to participate in the research program prior to undergoing an audiotaped, semistructured memory interview that enabled us to place them into one of the four groups. Reported sexual abuse ranged from fondling to rape. Susan A. Clancy also conducted psychiatric diagnostic interviews to ascertain the presence of current mental disorders (e.g., posttraumatic stress disorder [PTSD], major depression; First, Spitzer, Gibbon, & Williams, 1996; Foa & Tolin, 2000).

Participants enrolled in our research program visited our laboratory three times to participate in interviews and different experimental tasks. Individuals received $25 per visit as an honorarium. The paradigm in the present study was described to participants as a “letter-height judgment task.” Heidi M. Barrett and Holly A. Parker tested participants. Although these two experimenters were aware of the hypotheses, we endeavored to keep them blind as to participant’s group (e.g., repressed memory group).
However, it was impossible to ensure that participants did not make inadvertent comments to the experimenters, thereby revealing their group.

**Continuous memory group.** The continuous memory group consisted of 72 adults (54 women, 18 men) who said they had never forgotten their CSA. Sixteen furnished the name of an informant who could corroborate the abuse, and another participant provided abuse-related entries from a diary that dated from when the abuse had occurred. Susan A. Clancy spoke to all informants, who corroborated the accuracy of the participants’ reports. Six met criteria for major depression, 28 met criteria for abuse-related PTSD, and 5 met criteria for PTSD unrelated to abuse.

**Recovered memory group.** The recovered memory group consisted of 26 adults (17 women, 11 men) who reported having recovered long-forgotten memories of CSA. Individuals in this group reported at least one autobiographical memory of CSA that came to mind after many years of having not thought about the abuse; on average, 21.5 years had elapsed. Although about 75% sought psychotherapy at some point in their lives, only two participants had recalled their abuse during the course of psychotherapy (Clancy & McNally, 2003). No participant was either able or willing to furnish corroboration of the abuse. One met criteria for major depression, eight met criteria for abuse-related PTSD, and two met criteria for PTSD unrelated to abuse.

We use the term recovered memory participant in a purely descriptive sense to refer to someone who reports thinking about an alleged CSA event that he or she had not thought about in many years. Lacking corroboration of these reported events, we can neither endorse them as true nor claim they are false. Moreover, we use the term recovered memory without making any commitment about the mechanisms that might explain why the CSA memory (assuming it is true) had not surfaced earlier (e.g., lack of retrieval cues; voluntary suppression; dissociation).

**Repessed memory group.** The repessed memory group consisted of 38 adults (32 women, 6 men) who believed that they had been sexually abused during childhood but who had no autobiographical memories of their molestation. They inferred their abuse from various sources, such as visual images interpreted as “flashbacks,” dreams interpreted as abuse memories, or from sexual problems or other distress.

We use the term repessed memory participant in a purely descriptive sense to refer to someone who believes that he or she harbors inaccessible memories of CSA. Lacking any memories to corroborate, we cannot say whether any of these individuals had been abused. Moreover, our use of the term repessed does not imply any commitment to a mechanism of inaccessibility; it merely describes the phenomenology of individuals in this group. Four met criteria for major depression, and three met criteria for PTSD unrelated to CSA.

**Comparison group.** The comparison group consisted of 26 adults (16 women, 10 men) who reported no history of CSA. None met criteria for either major depression or PTSD.

### Table 1

**Means for Demographic and Psychometric Data**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Comparison</th>
<th>Continuous</th>
<th>Recovered</th>
<th>Repressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age (years)</td>
<td>38.3</td>
<td>14.2</td>
<td>37.9</td>
<td>10.6</td>
</tr>
<tr>
<td>Education (years)</td>
<td>15.3</td>
<td>1.7</td>
<td>14.8</td>
<td>2.3</td>
</tr>
<tr>
<td>DES score</td>
<td>8.2</td>
<td>7.7</td>
<td>17.4</td>
<td>13.6</td>
</tr>
<tr>
<td>Absorption score</td>
<td>14.6</td>
<td>7.5</td>
<td>18.1</td>
<td>7.1</td>
</tr>
<tr>
<td>VVIQ score</td>
<td>65.6</td>
<td>8.5</td>
<td>62.4</td>
<td>8.9</td>
</tr>
</tbody>
</table>

*Note.* Because of missing data, degrees of freedom vary. DES = Dissociative Experiences Scale; VVIQ = Vividness of Visual Imagery Questionnaire.

## Questionnaires

Participants completed a battery of questionnaires that included the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973), the DES (Bernstein & Putnam, 1986), and the Absorption Scale (Tellegen & Atkinson, 1974). Finally, we recorded age and years of education.

The means and standard deviations for these variables are shown in Table 1. Analyses of variance revealed no reliable (ps > .05) differences among the groups in terms of age, years of education, absorption, or imagery ability. A significant difference in dissociation symptoms was attributed to the comparison group scoring lower than the other groups as confirmed by post hoc contrasts (least-significant differences method).

## Results

Standard formulas for calculating $d'$ are undefined for hit rates of 1.0 and false alarm rates of 0 because the corresponding $z$ scores for these values are infinite. To prevent these problems, we followed Snodgrass and Corwin’s (1988) recommendations and computed a corrected hit rate by adding 0.5 to the number of words correctly affirmed by the participant as having been seen and dividing this value by 13 (i.e., the number of words that had been presented on cards plus 1).\(^1\) We computed a corrected false alarm rate by adding 0.5 to the number of words incorrectly affirmed by the participant as having been seen and dividing this value by 13 (i.e., the number of words that had been presented on audiotape plus 1).

The assumptions of signal detection theory require the statistical independence of the measures of sensitivity and criterion (response bias). Accordingly, we used C rather than $\beta$ as the measure of a participant’s criterion for affirming having seen a word (Snodgrass & Corwin, 1988). Therefore, our formulas for sensitivity and criterion were $d' = z_{FA} - z_{H}$ and $C = 0.5(z_{FA} + z_{H})$ (Snodgrass & Corwin, 1988).

We used one-tailed focused contrasts to test our specific predictions, and we computed effect size $r$ for each contrast. Focused

---

\(^1\) Due to experimenter oversight, the word *fish* was used as an example when discussing the procedure with several participants. Because *fish* appeared as a word in some later stimulus sequences, we eliminated this word from the analyses, thereby resulting in a denominator of 11, not 12, in the calculation of hit rates and false alarms. Also, one participant in the continuous memory group did not complete the second trial.
contrast analysis is a more powerful method for testing specific hypotheses than are omnibus analyses of variance (Rosenthal & Rosnow, 1985).2

Sensitivity ($d'$)

To the extent that individuals experience difficulty discriminating products of perception from products of imagination, they should be prone to confuse fantasies with memories. Therefore, skeptics of repression predict that individuals reporting repressed and recovered memories of CSA should have lower $d'$ scores than should individuals reporting continuous memories of CSA or those reporting no history of CSA. One-tailed contrasts nearly confirmed this hypothesis for the first trial block, $t(160) = 1.49, p = .07$, effect size $r = .12$, and confirmed it for the second trial block, $t(159) = 2.47, p = .01$, effect size $r = .19$ (see Figure 1).

Criterion (C)

If the repressed and recovered memory groups have a low criterion (response bias) for affirming having seen versus having imagined a word relative to the continuous memory and comparison groups, then the C scores for the first two groups should be lower than for the second two groups. One-tailed contrasts failed to confirm this hypothesis for either the first trial block, $t(160) = -0.99, p = .16$, effect size $r = .08$, or for the second trial block, $t(159) = -1.67, p = .05$, effect size $r = .13$. Indeed, the negative contrast is indicate a pattern opposite to prediction (see Figure 2).

Change in Sensitivity ($d'$)

Finally, as evinced by Figure 1, all groups improved their sensitivity after receiving foreknowledge about the memory test prior to the second trial block. Even the repressed and recovered memory groups exhibited greater $d'$ values for the second block relative to the first.

As an additional test of the false memory/reality-monitoring hypothesis, we contrasted the $d'$ scores of the group whose memories were most likely genuine (continuous memory group) with those of the group whose “memories” were least likely to be genuine (repressed memory group). A one-tailed contrast confirmed that sensitivity for discriminating percepts from images was, indeed, greater for the continuous memory group than for the repressed memory group for the first trial block, $t(108) = 1.84, p = .01$, effect size $r = .17$, but not for the second one, $t(107) = 0.60, p = .28$, effect size $r = .06$. That is, in the absence of foreknowledge about the memory test following the second block, the participants reporting repressed memories were less able to discriminate percepts from images relative to participants who had never forgotten their abuse.

Predictors of Sensitivity ($d'$)

Finally, inconsistent with prediction, imagery ability (VVIQ), absorption, and dissociation tendencies (DES) were not signifi-

---

2 Because some readers may be unfamiliar with focused contrast analyses, we conducted a traditional omnibus analysis of variance (ANOVA):

An ANOVA on the C scores revealed the following effects: group, $F(3, 159) = 2.79, p = .04$; trial block, $F(1, 159) = 47.53, p = .01$; Group × Trial Block, $F(3, 159) = 0.67, p = .57$. According to post hoc contrasts, the comparison group had a lower C than the continuous memory and repressed memory groups. The trial block effect was attributable to all groups, lowering their C from the first to the second trial blocks. That is, the comparison group had the most liberal criterion for affirming that a word had been seen.

A 4 (group: comparison, continuous, recovered, repressed) × 2 (trial block: first, second) ANOVA on the $d'$ scores revealed the following effects: group, $F(3, 159) = 2.13, p = .10$; trial block, $F(1, 159) = 99.97, p < .01$; Group × Trial Block, $F(3, 159) = 3.03, p = .03$. The significant effect of trial block confirms the expected improvement in $d'$ evident in all groups. As revealed by post hoc contrasts (least-significant differences method), the interaction was attributable to the comparison group having a significantly higher $d'$ than the repressed memory group on the first block and a significantly higher $d'$ than the recovered memory group on the second block.
significantly (inversely) correlated with $d'$. Pearson correlations between $d'$ (first block) and these variables were trivial, nonsignificant ($p > .05$), and sometimes opposite to prediction: VVIQ, $r(155) = .08$; absorption, $r(156) = .12$; DES, $r(158) = .02$. The same held for the second block: VVIQ, $r(154) = -.01$; absorption, $r(155) = -.04$; DES, $r(157) = -.13$.

Discussion

Some psychologists believe that people who report repressed and recovered memories of traumatic events are often mistaken. They argue that what gets adduced as “memories” of long-buried trauma are not, in fact, cognitive representations of genuine events—memories arising from percepts. Dream fragments, repetitive disturbing images, “body memories,” and so forth, they propose, are usually the products of imagination, not perception. These arguments imply that people prone to report repressed and recovered memories of CSA are characterized by a reality-monitoring deficit that makes it difficult for them to discriminate between products of perception and products of imagination.

Our results provided inconsistent support for this perspective. Consistent with the hypothesis regarding a reality-monitoring deficit, the recovered and repressed memory groups had a marginally lower $d'$ than did the comparison and continuous memory groups for the first trial block and a significantly lower $d'$ for the second trial block. But as evident from Figure 1, the near-significant contrast for the first trial block was driven mainly by the difference between the comparison and repressed memory group; in fact, the recovered memory group tended to show better performance than the continuous memory group. Provision of foreknowledge about the memory test resulted in improved performance across the board, and now $d'$ scores were significantly lower in the repressed and recovered memory groups than in the comparison and continuous memory groups.

It is unclear why none of the individual difference variables predicted $d'$ scores. Each of these variables has predicted memory errors in previous research. For example, the DES has predicted false recognition in DRM studies (e.g., Clancy et al., 2000; Winograd, Peluso, & Glover, 1998) and false autobiographical memory formation in some studies (e.g., Hyman & Billings, 1998), but not in others (e.g., Spanos, Burgess, Burgess, Samuels, & Blois, 1999). The Absorption Scale has predicted false recognition and recall in DRM research (e.g., Clancy et al., 2002), as has the VVIQ (Winograd et al., 1998). However, the procedures in the aforementioned experiments differed from those in the present study. Moreover, it is possible that a purer measure of fantasy proneness, other than the Absorption Scale or the DES, might have correlated (negatively) with $d'$ scores (for a review of measures, see Eisen & Lynn, 2001).

Although there was some evidence of diminished sensitivity for distinguishing the products of perception from those of imagination in the repressed and recovered memory groups, there was no evidence of a response bias favoring affirmation of having seen rather than having imagined stimuli. If anything, the groups reporting abuse histories were more conservative than the comparison group in claiming to have seen items.

Further research along two lines is warranted. First, it would be useful to investigate reality monitoring with paradigms other than H. D. Brown et al.’s (1994). Although Brown et al.’s procedure requires participants to produce word stimuli in imagination to answer correctly questions about the letter heights, it is less well studied than procedures developed by Johnson and her colleagues. Accordingly, investigators should consider using these older methods for studying possible reality-monitoring deficits in abuse-reporting groups. Suffice it to say, signal detection methods in either a Brown et al. or a Johnson paradigm are capable of

---

1. Although imagination is required to perform the letter height judgment task correctly, we unfortunately did not track accuracy of these judgments as a manipulation check.
distinguishing two bases for memory errors—sensitivity and response bias—and therefore merit application whenever possible.

Second, although tasks such as H. D. Brown et al.’s (1994) enable assessment of sensitivity for distinguishing percepts from images in general, they do not reveal whether capacity to distinguish reality from fantasy varies as a function of the emotional significance of the material. Researchers should consider testing whether people reporting repressed or recovered memories of CSA exhibit reality-monitoring deficits for material possessing personal emotional significance.

References


Received July 31, 2003
Revision received May 14, 2004
Accepted May 20, 2004