Autobiographical Memory Specificity After Manipulating Retrieval Cues in Adults Reporting Childhood Sexual Abuse

Beatrijs J. A. Hauer  
Maastricht University

Elke Geraerts and Harald Merckelbach  
Maastricht University

Ineke Wessel  
University of Groningen

Tim Dalgleish  
Medical Research Council Cognition and Brain Sciences Unit

Traumatized samples have relative difficulty in generating specific autobiographical memories on a cue word task, compared to nonexposed controls. Simultaneously, trauma is associated with highly specific intrusive trauma memories in day-to-day life. Possibly, day-to-day intrusions and memories generated to cue words rely on different retrieval processes. With the former dependent on close associations between retrieval cues and specific memory representations (direct retrieval), and the latter on iterative retrieval cycles through a hierarchical memory system (generative retrieval). This study investigated this distinction using two versions of the cue word task, designed to promote generative and direct retrieval, respectively, in participants with or without a history of child sexual abuse (CSA). The data demonstrated that CSA participants were less specific than nonabused controls to generative retrieval cues, but this difference disappeared with direct retrieval cues. This interaction was stronger in CSA participants with relatively greater posttraumatic stress and remained significant when participants with past or current major depressive disorder were excluded and also when only those participants with corroborated CSA were included.

**Keywords:** autobiographical memory specificity, childhood abuse, depression, intrusions, trauma

Numerous studies suggest that individuals with emotional disorders have problems accessing specific aspects of their personal pasts. For example, Williams and Broadbent (1986) asked depressed suicide attempters and healthy controls to generate specific autobiographical memories to a series of emotion-related cue words (i.e., the Autobiographical Memory Test [AMT]). They found that the parasuicide group was far more likely to produce overgeneral responses (e.g., to the word party, an overgeneral response might be “I have never enjoyed parties”) than to reply with memories of specific experiences (e.g., “I enjoyed Jane’s party last Saturday”). Such difficulty in producing specific memories on the AMT has since been demonstrated in various clinical groups (see Williams et al., 2007, for a review), with the majority of studies focusing on depressed states (e.g., Brittlebank, Scott, Williams, & Ferrier, 1993; Kuyken & Dalgleish, 1995; Wessel, Meeren, Peeters, Arntz, & Merckelbach, 2001). Extending this research into the domain of trauma, Kuyken and Brewin (1995) showed that depressed women reporting childhood abuse showed reduced specificity relative to depressed women reporting no abuse. This relationship between a history of trauma and reduced memory specificity has since been replicated numerous times and following various kinds of traumatic events (see Moore & Zoellner, 2007, for a review).

The Affect Regulation Hypothesis

Theoretical conceptualizations of reduced autobiographical memory specificity have proposed various causal processes (Williams, 2006; Williams et al., 2007) grounded in mainstream cognitive models of autobiographical recall (e.g., the self-memory system; Conway & Pleydell-Pearce, 2000). In the context of trauma, a predominant and compelling view of the reduced specificity effect has been the affect regulation hypothesis (Williams, Stiles, & Shapiro, 1999). The proposal here is that trauma-exposed individuals experiencing trauma-related distress seek to avoid remembering the specific details of their traumatic experiences in an attempt to reduce that distress. This avoidant memory style, it is proposed, then generalizes beyond the recollection of the trauma memories until it characterizes much broader domains of autobiographical remembering, resulting in reduced specificity on tests such as the AMT, which do not explicitly target trauma memories.

The affect regulation hypothesis thus provides a plausible account of reduced memory specificity on the AMT in distressed
trauma survivors. However, it fails to account for why those survivors continue, in many instances, to report frequent intrusive and very specific memories of their trauma in their daily lives (Brewin, Christodoulides, & Hutchinson, 1996; McNally, 2003) which, at their most disabling, form the core of psychiatric conditions such as posttraumatic stress disorder (PTSD; American Psychiatric Association, 1994). In other words, why are the same populations that routinely show relatively reduced memory specificity on the AMT nevertheless characterized by unwanted memories of specific autobiographical trauma-related material?

One possibility is that day-to-day trauma-related intrusions, on the one hand, and reduced specificity on the AMT, on the other hand, are the products of different kinds of memory retrieval processes (Brewin, Dalgleish, & Joseph, 1996; Conway & Pleydell-Pearce, 2000; Williams et al., 2007). For example, it has been suggested that the reduced specificity effect on the AMT is a result of disruption in the effortful “generative” retrieval of memories within a hierarchically organized memory architecture such as the self-memory system referred to above (Conway & Pleydell-Pearce, 2000; Williams et al., 2007). Generative retrieval within this model relies upon top-down memory search processes involving the use of conceptual representations to form the basis for the search. Once search criteria are specified, general information (referred to lifetime periods, e.g., “when I lived in Cambridge” or to particular classes of events, e.g., “my holidays in France”) at the top of the hierarchy is rapidly activated. Activation then spreads through the autobiographical knowledge base from such general event representations to representations of specific autobiographical memories recorded at the bottom of the hierarchy. Within this theoretical context, reduced memory specificity results from premature truncation of the generative search for specific memories, leading to the retrieval of general autobiographical summaries from higher up in the hierarchy. In the case of trauma, then, this truncated search would represent the generic avoidance tendency that is at the heart of the affect regulation hypothesis.

In contrast, it has been proposed that intrusive recollections of specific trauma memories in daily life result from a process of “direct” autobiographical memory retrieval (Brewin, Dalgleish, & Joseph, 1996; Conway & Pleydell-Pearce, 2000; Conway, Singer, & Tagini, 2004; Williams et al., 2007; see also Burgess & Shallice, 1996). Direct retrieval refers to situations in which an internal or environmental cue produces immediate activation of a specific memory, thus bypassing any generative search. It is important to note that this bypassing of the generative search process ensures that direct retrieval of memories, including trauma memories, is unaffected by any premature truncation of generative retrieval occurring in the service of affect regulation.

In summary, the theoretical proposal is that the affect regulation hypothesis is operationalized by way of truncated generative autobiographical memory retrieval in distressed trauma survivors (Williams et al., 2007), leading to difficulty accessing any specific material by this means. In contrast, retrieval of any specific memories by way of direct means, including retrieval of trauma memories, remains unaffected by such truncation of generative search processes.

The principal aim of the present study was to examine this putative distinction in the laboratory by contrasting performance on the standard version of the AMT (proposed to broadly elicit generatively retrieved responses) with performance on an AMT constructed to broadly elicit directly retrieved responses. In this way, we can examine whether the usual reduced memory specificity effect on the standard AMT in distressed trauma survivors, relative to controls, is significantly reduced or attenuated on an AMT designed to promote relatively direct retrieval processes. It is important to stress that these versions of the AMT do not need to focus on retrieval of trauma memories because the theoretical arguments outlined above propose that all generative retrieval will be relatively compromised and that all direct retrieval will be relatively spared in distressed trauma survivors.

To construct an AMT designed to promote direct retrieval, we considered existing theoretical proposals about memories that are putatively directly retrieved (Brewin, Dalgleish, & Joseph, 1996; Conway & Pleydell-Pearce, 2000). One potential reason for direct retrieval is the presence of an extensive repertoire of direct associations between cues in the internal and external environment and specific representations in memory. Such associations could have become established as a function of the frequency with which memories have been accessed in response to certain cues and thus represent the relative automatization of memory access (Moors & de Houwer, 2006). Another possibility that is not mutually exclusive is that direct retrieval is more likely for memories stored as image-based mental representations (e.g., Brewin, Dalgleish, & Joseph, 1996; Conway & Pleydell-Pearce, 2000). Such memories would be directly accessed by cues or reminders that either contain or are strongly associated with image-based information that overlaps with, or is highly related to, the content of the representations.

These potential processes of automaticity and of facilitated specific memory retrieval using highly imageable cues were investigated by Williams, Healy, and Ellis (1999). These authors showed that manipulating the imageability of the cue words on the AMT led to marked differences in the proportions of specific memories recalled, with concrete, highly imageable cues leading to significantly greater specificity, consistent with the recruitment of image-based processes during recollection. Furthermore, retrieval of specific memories to these highly imageable cues was relatively immune to the effects of mental load, thus suggesting a degree of automatization (Williams, Chan, Barnhofer, Eade & Healy, 2006, Study 3).

On the basis of these theoretical ideas and accompanying data, we opted in the present study for an AMT methodology with cues varying in imageability as a vehicle with which to try to operationalize the putative contrast between relatively direct and generative autobiographical recall in trauma-exposed individuals outlined above. In line with this, the first main hypothesis for the present study was that a trauma-exposed sample would show reduced memory specificity relative to controls not exposed to trauma on the standard version of the AMT, which in general uses cues of low imageability (Williams & Broadbent, 1986), but that this group difference would be significantly attenuated or eliminated on an AMT consisting of highly imageable, concrete cue words designed to promote direct retrieval.

A further key finding in the AMT and trauma literature is that, among trauma-exposed samples, those individuals with higher levels of posttraumatic stress or with PTSD show more marked reductions in specificity on the standard AMT than do those with less distress or no PTSD (see Moore & Zoellner, 2007). In fact, there is a well-replicated negative correlation between levels of intrusion of specific trauma memories in day-to-day life and levels
of memory specificity on the AMT (e.g., Kuyken & Brewin, 1995; Wessel, Merckelbach, & Dekkers, 2002).

The prototypical account of these relationships (e.g., Williams et al., 2007) is that higher distress (including more frequent intrusions) following trauma results in a greater emphasis on affect regulation by way of truncated generative retrieval, and thus reduced specificity on the standard AMT. However, ironically, the intrusions that are a key source of the distress in the first place are largely unaffected by these regulation attempts by virtue of the fact that they are habitually directly retrieved and, as we have sought to argue, direct retrieval is relatively unaffected by truncated generative search.

Our second hypothesis sought to draw upon these theoretical claims. We predicted that within our trauma-exposed group, those survivors with clinically significant levels of posttraumatic stress (according to established cutoffs on the Impact of Event Scale [IES]; Horowitz, Wilner, & Alvarez, 1979) would show a more marked reduction in memory specificity on the standard AMT in comparison with those survivors with low levels of symptoms. In contrast, on the version of the AMT designed to promote direct retrieval, we predicted that this group difference would be attenuated or eliminated, as there is no reason to suppose that levels of traumatic distress should significantly affect specificity of directly retrieved general autobiographical material.

To investigate the hypotheses outlined above, we chose to recruit a sample of adults reporting a history of childhood sexual abuse (CSA) along with current intrusive symptomatology. This seemed an appropriate population because adult participants abused as children were the initial research focus in the trauma AMT literature (Kuyken & Brewin, 1995), and the reduced specificity finding on standard versions of the AMT has proven to be robust in this population (Moore & Zoellner, 2007). Furthermore, in theoretical terms it has been argued that childhood represents a critical period when individuals are highly susceptible to developing this particular memory style as a form of functional avoidance (Williams, Stiles, & Shapiro, 1999).

The recruitment of a CSA sample allowed us to address a supplementary hypothesis emerging out of the current AMT literature (Williams et al., 2007). Williams et al. noted that a significant limitation of the AMT findings with abused individuals to date (e.g., Kuyken & Brewin, 1995) was that in no instance was corroboration of the abuse reported. Consequently, our supplementary hypothesis (Hypothesis 3) was that those among our CSA participants for whom their history of abuse was corroborated would show reduced specificity relative to our non-CSA controls on the standard version of the AMT. Indeed, it seems reasonable to presume that the interaction effect specified in Hypothesis 1 should also be present across these two groups.

Reduced Memory Specificity and Depression History

To our knowledge, no studies describing reduced autobiographical memory specificity effects in traumatized samples have controlled for both past and current clinical depression. There are some studies that have controlled for current major depressive disorder (MDD; e.g., McNally et al., 2006; Schönfeld & Ehlers, 2006), but in these studies past MDD was not taken into account. As MDD is a common outcome following trauma (e.g., Shalev et al., 1998), this is not a trivial problem with the existing data in the literature, as it means that in all of the studies to date, reporting reduced specificity associated with trauma, history of MDD and trauma are confounded. As MDD history has been found to be associated with reduced memory specificity, even in the absence of a current MDD diagnosis and even when self-reported depressive symptoms have been controlled for (e.g., Mackinger, Pachinger, Leibetseder, & Fortacek, 2000), this leaves open the possibility that any apparently trauma-related effects on the AMT are somehow a function of group differences in MDD history that have not been adequately assessed (see the Methodological Limitations section in Williams et al.’s [2007] review of these findings).

In the present study we therefore sought to compare a subsample of the individuals with CSA who had never been clinically depressed with a subsample of non-CSA controls who had also never been clinically depressed to resolve this confound. Our fourth hypothesis was consequently that the never-depressed CSA group would show reduced memory specificity relative to the never-depressed non-CSA controls on the standard AMT after also covarying current levels of self-reported depressed mood. Indeed, we presumed that the interaction effect specified in Hypothesis 1 should also be present across these never-depressed subgroups. If this fourth hypothesis is supported by the data, this would indicate for the first time that trauma is significantly related to reduced specificity, independent of both current and past MDD.

Method

Participants

CSA participants (n = 70; 57 women; mean age = 41.37 years, SD = 11.59) were recruited through advertisements in provincial newspapers inviting participants to come to the laboratory if they had a history of CSA, of which they reported a continuous memory (i.e., participants reported that they had always remembered their CSA experiences). CSA was defined as sexual contact ranging from sexually related fondling to penetrative acts before the age of 12 (McNally, Clancy, Barrett, & Parker, 2004, 2005). In a standard semistructured interview (McNally et al., 2004, 2005), participants reporting CSA were asked a number of questions about the nature of their abuse (e.g., when it happened, what the duration was, and who the perpetrator was). All of the CSA participants reported current intrusive symptomatology.

Control participants (n = 63; 41 women; mean age = 39.06 years, SD = 12.57), group-matched with the CSA participants for age and education, were recruited through further advertisements describing research into personality and memory. Participants were only included in the control group if they denied having a history of CSA in the semistructured interview (McNally et al., 2004, 2005).

The Dutch version of the Structured Clinical Interview for the DSM–IV (SCID; Groenestijn, Akkerhuis, Kupka, Schneider, & Nolen, 1999) was used to assess MDD in both the CSA and the control groups. In the CSA group, 10 participants (14.3%) had a current major depressive episode. Thirty-one participants (44.3%) reported a history of MDD. Only 1 participant who had had a past MDD had a current major depressive episode. SCID–PTSD was also completed for those CSA participants with significant Criterion B symptoms of nightmares, flashbacks, and intrusive thoughts or images. There were 15 CSA participants (21.4%) who had
current PTSD with these Criterion B symptoms. Six of these participants were also diagnosed with a current MDD. Furthermore, 3 participants with current PTSD had formerly had an MDD. Abused participants were asked for corroboration of their abuse, either by other family members or by documentary evidence (e.g., letters, medical records). Thirty participants (42.9%) gave the name of a person who could corroborate the abuse (e.g., a sibling who was also abused), and confirmation was forthcoming in all cases. Two further participants (2.8%) provided a court document indicating that the perpetrator had been convicted for the abuse. In the control group, nobody was diagnosed with a current MDD. There were 8 participants (12.7%) with a previous MDD.

All participants received €50 and reimbursement of traveling expenses after their participation. This study was approved by the standing ethical committee of Maastricht University.

**Materials**

**AMT.** Two versions of the AMT (Williams & Broadbent, 1986) were used, each containing 10 words, 5 positive and 5 negative (available from Beatris J. A. Hauer). A Dutch translation (Wessel et al., 2001) of the original version of the AMT (the standard AMT) from Williams and Broadbent (1986) was compared to a version (the concrete AMT) comprising highly imageable, concrete cue words (e.g., funeral, cuddle). Words for the concrete AMT were selected on the basis of a pilot study (N = 13) in which 25 positive and negative words were rated on 9-point scales assessing imageability, concreteness, and how readily participants felt the words would bring to mind specific memories. The scales ranged from 1 (not at all visually imageable; extremely abstract; does not easily elicit a specific memory) to 9 (extremely visually imageable; extremely concrete; extremely easily elicits specific memory). Similar ratings were made for the standard AMT words. On the basis of these ratings, 10 words for the concrete AMT were selected such that they were significantly more imageable (M = 7.85; SD = .51 vs. M = 5.06; SD = .51, respectively), t(12) = 7.45, p < .001; more concrete (M = 7.77; SD = .55 vs. M = 4.60; SD = 1.20, respectively), t(12) = 7.60, p < .001; and elicited specific personal memories more readily than the standard words (M = 7.55; SD = .88 vs. M = 6.14; SD = .76, respectively), t(12) = 6.71, p < .001.1

On each AMT, participants were instructed to recall to each cue a specific memory about a personally experienced event that happened at a particular time and place that lasted less than 1 day. Participants were asked to provide a different memory for each cue. The cue words were orally and visually presented with positive and negative words alternating. To familiarize participants with the procedure, a maximum of 10 neutral practice words (e.g., shop, car, trip, conversation) were given until participants recalled at least three specific memories in direct succession. Participants were required to respond with a specific memory within 60 s. If participants gave a nonspecific answer before the 60 s, they were prompted one time with “Could you be more specific?” Once the time limit was reached, the experimenter presented the next cue word. The experimenter transcribed the exact response.

The first responses given to the cue words on the AMT were coded by an independent rater. Each response was coded as either a specific memory (i.e., referring to an event at a particular time and place, lasting less than a day), a general categoric memory (i.e., a group of similar events that had occurred on repeated occasions), a general extended memory (i.e., a single event lasting longer than 1 day), no memory (i.e., a coherent response that was not a memory; for example, a simple semantic association to the cue word), or an omission (i.e., no response). The total number of different specific memories retrieved to the 10 cue words for each version of the AMT was the primary dependent variable in the present study. However, for the first hypothesis we carried out a representative analysis with numbers of general categoric memories. A second rater scored all of the responses to the standard and concrete AMTs of a random sample of 20 participants (i.e., 400 memories) to assess interrater reliability. Interrater reliability for coding a memory as specific or nonspecific was good with a Kappa of .70.

**Beck Depression Inventory—Second edition (BDI–II).** The BDI–II consists of 21 items containing four statements that reflect depressive symptoms in increasing severity (scored 0–3; range, 0–63; Beck, Steer, & Brown, 1996; Van der Does, 2002). Internal reliability of the BDI–II in the present sample was excellent, Cronbach’s α = .94.

**IES.** The IES (Brom & Kleber, 1985; Horowitz et al., 1979) was administered to the CSA group to provide a measure of frequency of posttraumatic stress symptoms. The IES is a widely used instrument designed to measure frequency of intrusions and avoidance of memories related to the participant’s most negative stressful experience. This could be the CSA or another experience. Internal reliability of the IES in the present sample was excellent for the IES total scale, Cronbach’s α = .93, and also for the Intrusions and Avoidance subscales, α = .91 and α = .88, respectively.

**Procedure**

Participants first provided written, informed consent. They then completed the two versions of the AMT in a counterbalanced order, with a 15-min distractor task (an unrelated questionnaire) given in between the tests. After the two AMTs, the semistructured interview and the SCID were administered. Thereafter, participants filled in the BDI–II, the IES, and some other questionnaires not reported here. Finally, CSA participants were asked for any corroboration of their abuse.

**Results**

Table 1 presents the demographic and the self-report questionnaire data for the CSA and control groups. The groups did not differ in terms of age or education, although they did differ on gender ratio. The CSA group scored significantly higher than the controls on the BDI–II. Fifty-four participants (77%) in the CSA group filled in the IES with respect to their CSA experience. The remaining 16 participants cited a range of traumas including death

---

1 Although the present study is not concerned with valence effects, it is useful to note that on the basis of similar rating scales, the concrete words did not significantly differ from the standard words on positive valence (M = 7.33; SD = .58 vs. M = 7.38; SD = .75, respectively), t(12) < 1, p > .70, and negative valence (M = 3.85; SD = .55 vs. M = 3.51; SD = .88, respectively), t(12) = 1.91, p > .08, and on emotional intensity (M = 6.31; SD = 1.20 vs. M = 6.11; SD = .70, respectively), t(12) < 1, p > .50.
or illness, accident, and threat to beloved others or to the self. The IES scores of the CSA group were comparable to those reported in previous studies of trauma-exposed samples (de Decker, Hermans, Raes, & Eelen, 2003; Stokes, Dritschel, & Bekerian, 2004). Control participants filled in the IES with respect to their most negative stressful event. Thirty-two participants (51%) reported death or illness to the self or beloved others as their most negative event. The other control participants cited different traumas, including accident or threat to beloved others or to the self. Ten control participants completed the IES but did not report a reference event. For the whole sample, and for the CSA and control groups considered separately, the IES (including the subscales) and BDI were significantly intercorrelated, $r_{11022} = .26, p_{11021} = .05$.

### AMT Data

Tables 2 and 3 present the AMT data across the different CSA and control groups. Again, the various subsamples drawn from the CSA and control groups did not significantly differ regarding age and educational level. As for gender, the only subsamples that significantly differed in gender ratio were the CSA subsample, who filled in the IES with respect to their abuse and the control group, as well as the never-depressed CSA and never-depressed control groups. Yet, initial analyses including gender rendered no significant main effects or interactions involving this factor, all $F_{11021} < 1$. The same was true for initial analyses conducted with order of AMT presentation (standard– concrete vs. concrete– standard) as a factor. Therefore, all further analyses are reported without order and gender as factors.

**Psychometric equivalence of the two AMT tasks.** A concern when examining differential degrees of group differences across tasks is that any significant interactions may simply reflect differences in the discriminating power of the tasks at hand (e.g., Chapman & Chapman, 1973, 1978). We therefore calculated the true variance of the two versions of the AMT as an index of

### Table 1

**Demographic and Self-Report Data for Childhood Sexual Abuse (CSA) and Control Groups**

<table>
<thead>
<tr>
<th>Variable</th>
<th>CSA group ($n = 70$)</th>
<th>Control group ($n = 63$)</th>
<th>Test statistic</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age</td>
<td>41.37</td>
<td>39.06</td>
<td>$t_{131} = 1.10$</td>
<td>.27</td>
</tr>
<tr>
<td>% women</td>
<td>81.4</td>
<td>65.1</td>
<td>$\chi^2(1) = 4.57$</td>
<td>.03</td>
</tr>
<tr>
<td>Education*</td>
<td></td>
<td></td>
<td>$\chi^2(6) = 4.61$</td>
<td>.60</td>
</tr>
<tr>
<td>Self-reported symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI-II</td>
<td>12.51</td>
<td>4.46</td>
<td>$t_{130} = 4.81$</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>IES total</td>
<td>26.86</td>
<td>13.95</td>
<td>$t_{125} = 4.04$</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Injuries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IES</td>
<td>13.10</td>
<td>6.02</td>
<td>$t_{125} = 4.08$</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Note. BDI-II = Beck Depression Inventory—Second Edition (Beck, Steer, & Brown, 1996; Van der Does, 2002); IES = Impact of Event Scale (Brom & Kleber, 1985; Horowitz et al., 1979).*  
+Scores ranged from 1 = elementary school to 7 = academic degree.  
+b One participant in the CSA group did not fill in the BDI-II and the IES. In the control group, 5 participants did not fill in the IES. For this reason, the degrees of freedom are lower in the analyses in which these measures are used.

### Table 2

**Autobiographical Memory Test (AMT; Williams & Broadbent, 1986) Data Across Childhood Sexual Abuse (CSA) and Control Groups**

<table>
<thead>
<tr>
<th>Type of memory</th>
<th>CSA group ($n = 70$)</th>
<th>Control group ($n = 63$)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard AMT</td>
<td>Concrete AMT</td>
<td>Standard AMT</td>
<td>Concrete AMT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Specific</td>
<td>4.84</td>
<td>2.77</td>
<td>7.03</td>
<td>2.01</td>
<td>6.24</td>
<td>2.28</td>
<td>7.41</td>
<td>2.41</td>
</tr>
<tr>
<td>Specific same event*</td>
<td>.17</td>
<td>.45</td>
<td>.06</td>
<td>.29</td>
<td>.05</td>
<td>.21</td>
<td>.02</td>
<td>.13</td>
</tr>
<tr>
<td>General categoric</td>
<td>1.23</td>
<td>1.11</td>
<td>.59</td>
<td>.91</td>
<td>.40</td>
<td>.81</td>
<td>.31</td>
<td>.60</td>
</tr>
<tr>
<td>General extended</td>
<td>1.33</td>
<td>1.50</td>
<td>.61</td>
<td>.79</td>
<td>.71</td>
<td>.76</td>
<td>.80</td>
<td>.87</td>
</tr>
<tr>
<td>No memory</td>
<td>1.80</td>
<td>1.87</td>
<td>1.34</td>
<td>1.54</td>
<td>1.73</td>
<td>1.66</td>
<td>1.36</td>
<td>1.77</td>
</tr>
<tr>
<td>Omission</td>
<td>.63</td>
<td>1.05</td>
<td>.37</td>
<td>.85</td>
<td>.78</td>
<td>1.12</td>
<td>.11</td>
<td>.31</td>
</tr>
</tbody>
</table>

*a Specific same event = when a particular specific memory was generated two or more times to different cue words.
Table 3
Specific and General Categoric Autobiographical Memory Data on the Standard and Concrete Autobiographical Memory Tests (AMT: Williams & Broadbent, 1986) in Subsamples of the Child Sexual Abuse (CSA) and Control Groups

<table>
<thead>
<tr>
<th>Subsample</th>
<th>Standard AMT</th>
<th></th>
<th>Concrete AMT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific</td>
<td>Categoric</td>
<td>Specific</td>
<td>Categoric</td>
</tr>
<tr>
<td></td>
<td>M     SD</td>
<td>M     SD</td>
<td>M     SD</td>
<td>M     SD</td>
</tr>
<tr>
<td>Control-ND (n = 55)</td>
<td>6.33</td>
<td>.25</td>
<td>.40</td>
<td>.81</td>
</tr>
<tr>
<td>CSA filled in IES with respect to abuse (n = 54)</td>
<td>4.57</td>
<td>.27</td>
<td>1.24</td>
<td>.10</td>
</tr>
<tr>
<td>CSA-IES high (n = 23)</td>
<td>3.65</td>
<td>1.71</td>
<td>1.43</td>
<td>.90</td>
</tr>
<tr>
<td>CSA-IES low (n = 26)</td>
<td>5.42</td>
<td>.90</td>
<td>1.04</td>
<td>1.11</td>
</tr>
<tr>
<td>Control-IES low (n = 38)</td>
<td>6.63</td>
<td>2.48</td>
<td>.50</td>
<td>.89</td>
</tr>
<tr>
<td>CSA-ND (n = 28)</td>
<td>4.78</td>
<td>2.79</td>
<td>1.14</td>
<td>1.18</td>
</tr>
<tr>
<td>CSA-C (n = 32)</td>
<td>4.50</td>
<td>2.64</td>
<td>1.38</td>
<td>1.18</td>
</tr>
</tbody>
</table>

Note. Control-ND = subsample of never-depressed control participants; IES = Impact of Event Scale (Broun & Kleber, 1985; Horowitz et al., 1979); CSA-IES high = subsample of CSA participants with IES scores of >35; CSA-IES low = subsample of CSA participants with IES scores of <19; Control-IES low = subsample of control participants with IES scores of <19; CSA-ND = subsample of never-depressed CSA participants; CSA-C = subsample of CSA participants who had corroboration for their abuse.

Hypothesis 1: CSA versus control groups. Table 2 presents the AMT data across the CSA and control groups. A 2 (group: CSA vs. control) × 2 (AMT: standard vs. concrete) analysis of variance (ANOVA) with numbers of specific memories as the dependent variable was performed to investigate our first hypothesis that the CSA group would show reduced memory specificity on the standard AMT in comparison with controls, but that this group difference would be significantly attenuated or eliminated on the concrete AMT. There were significant main effects of AMT, F(1, 131) = 85.20, p < .001, and group, F(1, 131) = 6.25, p = .01, on the standard AMT. These were qualified by a significant Group × AMT interaction, F(1, 131) = 7.71, p = .006, .06. Independent samples t tests demonstrated that the groups differed significantly on the standard AMT, t(131) = 4.57, p < .001, d = .85, but not on the concrete AMT, t(131) = 1.69, p = .09. Again, including the BDI–II in the analyses did not change the results. For the remainder of the analyses, we report only those findings involving number.

discriminating power (Chapman & Chapman, 1973). True variance can be estimated by calculating the product of the internal reliability and the observed variance for each measure from a sample of healthy control participants. Using this method, the discriminating powers of the standard AMT and the concrete AMT, on the basis of the scores of the controls (n = 63) in the present study, were 3.32 and 2.65, respectively. It is interesting to note that the internal reliabilities for both versions of the AMT were virtually indistinguishable, Cronbach’s α = .64. For this reason we could statistically compare the two measures of true variance by simply comparing the observed variance using an F ratio test. On this basis, the two measures did not significantly differ from each other, p = .37. Thus, we were able to conclude that there was no support for the two AMTs being different from each other in terms of estimated discriminating power (e.g., Chapman & Chapman, 1973, 1978), thus giving us confidence that any significant interaction effects represented differential group differences that were unlikely simply to be psychometric epiphenomena.

2 Some authors have argued that an analysis of covariance (ANCOVA) is often misused when dealing with covariates that may be related to the group factor when group membership is not random (see Miller & Chapman, 2001). This is potentially relevant in the present study because it is plausible that CSA participants are more depressed than control participants for reasons that are directly associated with their history of abuse. In other words, membership in the CSA group might not be independent of being somewhat depressed. This means that covarying depressive symptoms from the group variable in the present analysis would only allow us to draw conclusions about whatever is residual about membership in a CSA group, once the variance associated with depressed mood has been accounted for. However, as Miller and Chapman (2001) also argue, if one is less concerned about what the residual group variable might actually represent in an ANCOVA and is simply motivated to discover whether this residual variable still produces significant effects of interest following covariation, then an ANCOVA is justified. These were the circumstances of the present analyses, in which we simply wanted to reassure ourselves that the significance of the effects associated with the group variable was not entirely carried by group differences in depressed mood.
bers of specific memories as the dependent variable. However, the patterns of findings were comparable to those findings involving numbers of specific memories when categoric memories were used.

As a minority of the CSA group (n = 16) did not identify abuse as their index trauma on the IES, the above analyses were repeated with these participants excluded. Mean specificity and categoric data for the remaining CSA participants (n = 54) are presented in Table 3. The results of these analyses with and without the BDI–II as a covariate yielded the same patterns of significant results as those previously reported.

Hypothesis 2: CSA participants with and without significant PTSD symptomatology (vs. controls). To investigate our second hypothesis, we divided our CSA group on the basis of their IES scores into those falling within Horowitz et al.’s (1979) nonclinical range (using the cutoff of <19; CSA-IES low; n = 26), and those falling within the likely PTSD range (CSA-IES high; n = 23) identified in more recent studies (using the cutoff of 35; Neal et al., 1994; Wohlfarth, van den Brink, Winkel, & ter Smitten, 2003). These CSA-IES high and CSA-IES low groups did not differ significantly in age, t(47) < 1; education, χ²(6) = 3.50, p > .60; or gender ratio, χ²(1) = .05, p > .50.

We compared the performance of these CSA groups with that of control participants with IES scores <19 (controls-IES low; n = 38) on the two AMT tasks (see Table 3), with a 3 (group: CSA-IES high vs. CSA-IES low vs. controls-IES low) × 2 (AMT: standard vs. concrete) ANOVA with numbers of specific memories as the dependent measure. The results revealed significant effects of group, F(2, 84) = 5.80, p = .004, η² = .12, and AMT, F(1, 84) = 65.18, p < .001, η² = .44, and a significant Group × AMT interaction, F(2, 84) = 9.49, p < .001, η² = .18.

To deconstruct the interaction term, we examined the relevant patterns of paired comparisons using independent samples t tests. On the standard AMT, the CSA-IES high group was significantly less specific than the CSA-IES low group, t(45.78) = 2.44, p < .02, d = .72, but the two groups did not differ on the concrete AMT, t < 1, thus replicating the pattern of findings involving the whole CSA group versus controls reported above. Unsurprisingly, the CSA-IES high group also differed from the control-IES low group on the standard AMT, t(59) = 4.97, p < .001, d = 1.32, and again the groups did not differ on the concrete AMT, t < 1, thus reproducing the same pattern. Finally, the difference between the CSA-IES low and the control-IES low groups was approaching significance on the standard AMT, t(62) = 1.88, p = .07, but not on the concrete AMT, t < 1, reflecting a significant AMT × Group interaction between these two low IES groups, F(1, 62) = 5.88, p < .02, η² = .04.3

As for the whole study sample, within the CSA group, memory specificity was significantly negatively correlated with self-reported depression on the BDI–II, standard AMT, r(69) = −.25, p = .04; concrete AMT, r(69) = −.38, p = .001. Consequently, the above analyses were repeated with BDI–II scores as a covariate, producing an identical pattern of results.

We also examined the correlations between IES scores and numbers of specific memories, this time within the CSA group as a whole. Spearman’s correlations were used as the IES scores were not normally distributed. For the standard AMT, the number of specific memories was significantly negatively correlated with IES Avoidance scores, r(67) = −.26, p < .04, but not with IES Intrusion scores, r(68) = −.14, p = .25, although the latter were in the anticipated direction. The significant correlations involving IES Avoidance scores were no longer significant once BDI scores were covaried, p(r65) = −.15, ns. For the concrete AMT, IES scores did not significantly correlate with specificity: for the IES-Avoidance subscale, r(67) = −.16, p = .19; for the IES Intrusion subscale, r(69) = −.07, p = .57. These data broadly reflect the initial analyses using the IES cutoffs reported above.4

Hypothesis 3: Corroborated CSA group versus control group. Table 3 also presents the AMT data for CSA participants with corroboration for their abuse (CSA-C; n = 32). A 2 (group: CSA-C vs. controls) × 2 (AMT: standard vs. concrete) ANOVA with specific memories as the dependent variable revealed significant main effects of AMT, F(1, 93) = 93.00, p < .001, η² = .38, and of group, F(1, 93) = 5.39, p = .02, η² = .11. These were qualified by a significant Group × AMT interaction, F(1, 93) = 11.27, p = .001, η² = .05. Subsequent independent samples t tests demonstrated that the CSA-C group was less specific on the standard AMT than controls, t(93) = 3.33, p = .001, d = .70, but not on the concrete AMT, t(93) < 1, p > .60, supporting our hypothesis. Again, the analyses yielded the same pattern of results with the BDI–II score covared.

Hypothesis 4: Never-depressed CSA participants versus never-depressed controls. Table 3 shows the AMT data for never-depressed CSA participants (CSA-ND; n = 28) and for never-depressed controls (controls-ND; n = 55). To investigate our fourth hypothesis, we conducted a 2 (group: CSA-ND vs. control-ND) × 2 (AMT: standard vs. concrete) ANOVA with number of specific memories as the dependent measure. There was a main effect of AMT, F(1, 81) = 56.12, p < .001, η² = .35, and the main effect of group was approaching significance, F(1, 81) = 3.08, p = .08, η² = .03. The key AMT × Group interaction was also significant, F(1, 81) = 9.23, p = .003, η² = .10. Independent samples t tests again revealed that the CSA-ND individuals were less specific than the control-ND group on the standard AMT, t(81) = 2.71, p = .008, d = .61, but not on the concrete AMT, t(81) < 1, p > .80. These results therefore support our third hypothesis. As before, the relevant analyses remained significant with BDI–II scores included as a covariate.

Discussion

The first hypothesis of the present study was that participants reporting CSA would exhibit significantly reduced memory spec-

---

3 It is interesting to note that, as can be seen in Table 3, this was a crossover interaction. This provides further evidence that the differential performance across the two tasks is not simply a function of their psychometric properties (Chapman & Chapman, 1973).

4 As a final examination of the relationship between IES scores, BDI–II scores, and specific memories, we carried out regression analyses to examine the unique contributions of the scores on the two self-report measures (as predictors) to the memory scores (as dependent variables). For the standard AMT, the overall model was significant, F(2, 129) = 7.77, p < .001, accounting for 11% of the variance in memory scores, and the standardized betas and part correlations were as follows: for the IES, β = −.12, r_part = −.09, r(127) = 1.11, ns; for the BDI–II, β = −.24, r_part = −.18, r(127) = 2.11, p < .05. For the concrete AMT, the overall model was not significant, F(2, 129) = 2.34, ns, and accounted for only 4% of the variance in memory scores.
ificity in comparison with controls not reporting CSA on the standard version of the AMT, chosen to elicit relatively generative memory retrieval, but that this difference would be significantly attenuated or absent on a concrete-word version of the AMT, designed to promote relatively direct retrieval of memories (Conway & Pleydell-Pearce, 2000; Williams et al., 2007). The data supported this hypothesis, even after covarying levels of depressed mood on the BDI-II and even when analyses were restricted to those individuals reporting corroboration of their abuse. To our knowledge, this is the first demonstration that the reliable reduced memory specificity effect can be eliminated in a trauma-exposed population, relative to controls, by manipulating the content of the cue words on the AMT.

The second hypothesis was that a similar pattern would emerge when comparing those survivors in the CSA group with clinically significant levels of posttraumatic stress (as measured on the IES) with those survivors with minimal symptoms. Again the data supported our hypothesis, even when self-reported depression on the BDI-II was covaried. This pattern was also mirrored in correlational analyses using the IES Avoidance subscale in the CSA group. These data are in accordance with the findings from a study by Schönfeld and Ehlers (2006), published after we began the present research. Schönfeld and Ehlers compared groups of participants with and without PTSD on the numbers of general memories they retrieved on a standard cue word AMT and on a version in which the cues were affective pictures. They found that the participants with PTSD were significantly less specific in response to the pictures than to the words, with no significant difference for the participants without PTSD. However, in contrast to the pattern in the present study, Schönfeld and Ehlers found that their PTSD group remained more overgeneral than did the controls without PTSD, even on the picture AMT (S. Schönfeld, personal communication, December 1, 2006). Possibly, picture cues match less well with individuals’ memories than words cues because they are too specific and therefore still elicit generative retrieval.

The findings pertaining to the above hypotheses are potentially important in two ways. First, the fact that the same trauma-exposed participants were either less specific relative to controls or were not significantly different from controls, dependent on the cueing context (with comparable findings for high- vs. low-symptom trauma-exposed groups), is consistent with theoretical conceptualizations of autobiographical recollection in distressed trauma survivors (Williams et al., 2007) that propose relatively distinct retrieval processes that operate under different circumstances. One possible implication of the present data is that processes that disrupt relatively generative retrieval of memories, such as affect regulation (Williams, Stiles, & Shapiro, 1999), have little or no effect on relatively direct recall of the personal past. If, as we assume, intrusive reexperiencing of the trauma in trauma survivors is a route for relatively direct retrieved memories and such retrieval is promoted by using the concrete word version of the AMT, this argument would suggest that processes of affect regulation targeted to avoid distressing, specific autobiographical material are in fact missing the primary source of such distress in the form of these intrusive trauma memories.

Evidence in support of this possibility is provided by Golden, Dalgleish, and Mackintosh (2007), who showed that individuals who had experienced a traumatic bereavement and were suffering from complicated grief (CG) did not show the usual reduced specificity effect, relative to bereaved controls not suffering from CG, on a cue word task (the Biographical Memory Task; BMT) designed to probe memories from the lifetime of the deceased loved one (rather than the self). In contrast, they did show the usual reduced specificity effect on the standard AMT and on a BMT with reference to the life of a loved one who was still alive. As memories of the deceased form the content of day-to-day distressing intrusions in CG sufferers, these data provide evidence supporting the view that retrieval of such memories is relatively immune to putative processes of affect regulation that operate on the standard AMT.

The present data and those of Golden et al. (2007) do not necessarily mean that affect regulation in the form of reduced memory specificity is wholly dysfunctional for traumatized individuals (in particular for those with significant symptoms), simply that it is imperfect. For example, it may be that such “avoidance” processes still mitigate the trauma-related distress experienced by survivors by sufficiently reducing access to trauma-related material by way of relatively generative routes and/or by reducing the retrieval frequency of other emotive memories. This is a matter for future research. However, it is important to note that even if these putative affect regulation processes do effectively compromise relatively generative recall of the trauma, this in itself may be dysfunctional over a medium- to long-term period. For example, there is a growing body of research indicating that successful and repeated generative recall of traumatic events as a function of prolonged exposure techniques is broadly beneficial in terms of the amelioration of posttraumatic stress (e.g., Ehlers, Clark, Hacking, McManus, & Fennell, 2005; Pennebaker, 1995). Furthermore, reduced autobiographical specificity is related to poor problem solving strategies (e.g., Goddard, Dritschel, & Burton, 1996, 1997).

The finding that a significant reduced specificity effect on the standard AMT can be eliminated by selecting different cue words (and attenuated by using pictures; Schönfeld & Ehlers, 2006) also has a potentially important methodological implication. This finding highlights the fact that care needs to be taken in selection of cue materials for future AMT experiments to ensure that the appropriate memory retrieval context is invoked (see also Williams et al., 2006; Williams, Healy, & Ellis, 1999).

Our third hypothesis was that the difference in specificity on the standard AMT between CSA participants and controls would remain significant when we included only those participants from the CSA sample with corroborated abuse. The data supported this hypothesis and, to our knowledge, this represents the first demonstration of the standard AMT effect involving a corroborated abuse sample.

The final hypothesis of the present study was that the subsample of our CSA group with no current or past MDD would be less specific on the standard AMT than would the never-depressed, non-CSA controls, even after covarying current levels of depressed mood on the BDI-II. The data confirmed this hypothesis, and indeed the pattern of findings in support of Hypothesis 1 was also

5 It is important to note there was also no significant Group × Task interaction for numbers of specific memories in the Schönfeld and Ehlers (2006) data set and that the significance of the interaction involving general memories did not survive controlling for verbal IQ.
replicated across these subsamples. This gives us more confidence when interpreting the reduced memory specificity effect as it relates to trauma independent of MDD, in the service of making theoretical arguments within this literature.

It is important to note that we are not trying to argue that reduced specificity effects in traumatized populations (such as the current CSA group) are not influenced by depressed mood or other symptoms of depression. Indeed, in the present data, self-reported depressive symptoms on the BDI–II were negatively correlated with memory specificity, both across the sample as a whole and within the CSA group, on both concrete and standard versions of the task. This is perhaps not surprising given the high degree of intercorrelation between traumatic stress and depressive symptoms. Rather, what we have sought to do in the current analyses is demonstrate that there remains a significant relationship between AMT performance and trauma once the influence of depression is controlled for, both by selecting participants with no history of MDD and by covarying self-reported depressive symptoms on the BDI–II.

The present study is not without some limitations. The most significant one is that we included no independent way of assessing that the words in our concrete AMT invoked direct retrieval, as far as one can actually validate such theoretical constructs (see the introduction). It may therefore be the case that for some participants and/or for some cue words in the concrete AMT condition, the resulting retrieval processes were not direct (Conway & Pleydell-Pearce, 2000). However, to the extent that this was the case, it would have gone against the grain of our principal hypothesis (leading to reductions in memory specificity in the CSA group for these words on the concrete AMT). Consequently, this issue is perhaps less of a concern for the present study.

The fact that we did not assess direct versus generative retrieval explicitly in the present study of course means that we have no data that verify that the key difference between the two AMT tasks maps closely onto this theoretical distinction. It would be highly desirable for future studies to explicitly assess the extent to which retrieval to highly imageable word cues reflected proposed theoretical components of direct retrieval such as automaticity or access through image-based representations. Such studies could use manipulations of mental load, assess the sensory properties of the retrieved memories, or examine the time course of retrieval to different cues. Indeed, once the notions of direct versus generative retrieval can be deconstructed in terms of such other variables, it may be that the broader notions themselves prove redundant. This is a matter for further study.

A further limitation of the current study is that we did not carry out a complete assessment of trauma history in our control group, as our primary concern was to establish absence of self-reported CSA. It may be that some of the controls had experienced significant traumas that were not CSA with associated effects on their AMT performances. Again, however, this would have gone against the grain of our main hypothesis of a group difference on the standard AMT, relative to the concrete AMT, and so is of less concern for the present study.

A final limitation is the absence of reliability data for the SCID. Furthermore, the full set of PTSD questions for the SCID was only administered to those participants endorsing intrusion-related Criterion B symptoms. This therefore precluded formal comparisons between CSA participants with and CSA participants without PTSD. However, we were able to perform analogous comparisons with high- versus low-PTSD symptom groups using established cutoffs on the IES, and the results were in line with the study hypotheses.

In conclusion, the present study showed that the widely replicated finding of significantly reduced autobiographical memory specificity in trauma-exposed individuals, relative to controls, can be eliminated by the use of cue words designed to promote relatively direct autobiographical memory retrieval, with a similar pattern in high- versus low-symptom-level trauma survivors. The study also provided the first evidence that we know of that the reduced memory specificity effect associated with CSA is present independently of current and past MDD, and after controlling for current depressed mood.

References

Groenestijn, M. A. C., Akkerhuis, G. W., Kupka, R. W., Schneider, N., & van...


Received May 15, 2006
Revision received October 2, 2007
Accepted October 5, 2007