MEMORY, 2007, 15 (5), 495–507

Psychology Press
Taylor & Francis Group

A sentence completion procedure as an alternative to the Autobiographical Memory Test for assessing overgeneral memory in non-clinical populations

Filip Raes and Dirk Hermans
University of Leuven, Belgium

J. Mark G. Williams
University of Oxford, UK

Paul Eelen
University of Leuven, Belgium

Overgeneral memory (OGM) has been proposed as a vulnerability factor for depression (Williams et al., 2007) or depressive reactivity to stressful life-events (e.g., Gibbs & Rude, 2004). Traditionally, a cue word procedure known as the Autobiographical Memory Test (AMT; Williams & Broadbent, 1986) is used to assess OGM. Although frequently and validly used in clinical populations, there is evidence suggesting that the AMT is insufficiently sensitive to measure OGM in non-clinical groups. Study 1 evaluated the usefulness of a sentence completion method to assess OGM in non-clinical groups, as an alternative to the AMT. Participants were 197 students who completed the AMT, the Sentence Completion for Events from the Past Test (SCEPT), a depression measure, and visual analogue scales assessing ruminative thinking. Results showed that the mean proportion of overgeneral responses was markedly higher for the SCEPT than for the standard AMT. Also, overgeneral responding on the SCEPT was positively associated to depression scores and depressive rumination scores, whereas overgeneral responding on the AMT was not. Results suggest that the SCEPT, relative to the AMT, is a more sensitive instrument to measure OGM, at least in non-clinical populations. Study 2 further showed that this enhanced sensitivity is most likely due to the omission of the instruction to be specific rather than to the SCEPT's sentence completion format (as opposed to free recall to cue words).

It is widely agreed that *autobiographical memory* (AM) is essential to human functioning (for elaborate discussions of this point, see Conway & Pleydell-Pearce, 2000; Nelson & Fivush, 2004). ''AM'' has often been conceptualised as a subset of episodic memory. Alternatively, AM has also often been used synonymously with the term ''episodic memory''. We agree with recent authors (e.g., Conway, 1990, 2001, 2005) that both those conceptualisations of AM are misleading, and that episodic and autobiographical memory are best treated as two different, non-equivalent memory systems. Episodic memory contains personally experienced specific past events that lasted from minutes to hours. AM, however, is a more encompassing system, containing not only specific experience-near (episodic) self-related knowledge high in sensory-perceptual detail, but also more generic conceptual self-related material which is more semantic in nature. This view of AM as a

Address correspondence to: Filip Raes, Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium. E-mail: filip.raes@psy.kuleuven.be

We thank Sara Verellen for her assistance in the conduct of this study.

http://www.psypress.com/memory DOI:10.1080/09658210701390982

memory system containing personal knowledge ranging from specific episodic events to more generic semantic self-related information is reflected in, for example, Conway's framework of the Self-Memory System (SMS; Conway & Pleydell-Pearce, 2000; Conway, Singer & Tagini, 2004), and this view is also consistent with multiple-memory models (e.g., Johnson, 1992; see Williams et al., 2007).

When people are asked to retrieve a specific personal event, e.g., in response to cue words, such (strategic) retrieval most often reflects a staged process, beginning with the retrieval of more conceptual abstract (semantic-*generic*) self-related information moving on to more event-*specific* (episodic) self-related material high in sensory-perceptual detail (see Conway & Pleydell-Pearce, 2000; Williams et al., 2007). Interestingly, a recent line of study in the field of AM has focused on that very feature of personal memories; that is, the level of specificity (vs generality) with which such self-related knowledge is retrieved. More specifically, results obtained during the past 20 years have consistently shown that depressed people have more difficulty retrieving specific memories from AM than non-depressed people and generate relatively more personal information from the more semantic-generic level of autobiographical knowledge (see Williams et al., 2007, for a review).

In that (clinically oriented) research domain, the *Autobiographical Memory Test* (AMT; Williams & Broadbent, 1986) has most often been used. In the AMT respondents are instructed to retrieve a specific memory in response to cue words (e.g., *sad*, *happy*). A specific memory refers to a personal past experience that happened on a particular day and did not last longer than a single day (e.g., ''I had good fun at my brother's wedding two years ago''). As compared to non-depressed controls, depressed people generate relatively more overgeneral memories to cue words on the AMT (e.g., ''the times I go out for a drink'', rather than ''that Friday evening when we went for a drink at the Hopduvel in Ghent with some colleagues after a meeting'').

This particular memory phenomenon has been termed *overgeneral memory* (OGM) and refers to a tendency in some people, notably depressed ones, to retrieve past experiences in an overgeneral way; that is, either in the form of so-called overgeneral categoric memories or as overgeneral semantic associations. Categoric memories refer to summaries of similar events (e.g., ''every time I have to say goodbye or let go of something''). Overgeneral semantic associations refer to overgeneral (personal) semantic information (e.g., ''my mum'', ''my husband'', ''my home'').[1] According to recent theorising (e.g., Conway & Pleydell-Pearce, 2000; Conway et al., 2004; Williams et al., 2007), semantic associates are situated at an even higher level than ''categoric memories'' in the hierarchy of memory retrieval. One could thus say that these semantic associates constitute a sort of ''over-overgeneral'' response. As such, they are also—and maybe even more so than categoric memories—an indication of OGM (Williams, 2000).[2] Consider the following example. When a person responds ''home'' (i.e., a semantic associate) to the cue *happy*, this could be regarded as an *over*-overgeneral categoric memory. It collapses over categoric memories such as: ''moments with my children, playing in our garden'', ''when I'm with my husband, watching television in our living room'', ''when my husband and children join me for dinner in our kitchen'', ''enjoying the sun while reading a book in my rocking chair on the terrace'', ''moments with my cat Felix in the couch'', etc. In line with this, OGM will be defined here as the retrieval of overgeneral categoric *and* overgeneral personal semantic information. The inclusion of semantic associates in the categoric code is in line with recent theoretical models of AM (see above)

[1] Respondents sometimes retrieve overgeneral memories to cues that are not *categoric* but *extended* in that they refer to an extended period of time (lasting longer than a day). Examples of such overgeneral extended memories are ''when I was in high-school'', ''my time as a student in Leuven'', ''my holiday in Zimbabwe last year''. Subsequent studies following the seminal study of Williams and Broadbent (1986) made clear that overgeneral memory observed in clinical groups, such as in depressed and suicidal patients, is attributable to an increase in the *categoric* type of overgeneral memories (e.g., Barnhofer, de Jong-Meyer, Kleinpass, & Nikesch, 2002; Williams & Dritschel, 1992), but not the *extended* type.

[2] In fact, at the AMT Consensus Meeting (Williams, 2000), it was explicitly mentioned that semantic associates are indeed to be situated at a higher level in the hierarchy of overgenerality, and thus are to be regarded as sort of over-overgeneral memories. Since that meeting, it was recommended that if researchers wish to score the specificity of AMT responses on a numeric scale, semantic associates should be given the lowest score (i.e., ''1''), with categoric memories given a ''2'', extended memories a ''3'', and specific memories a ''4''. However, researchers have never really used this way of scoring AMT responses. By tradition, researchers mainly just look at the amount or proportion of categoric and specific responses, but not semantic associates.

as well as with original guidelines that resulted from an expert consensus meeting (see also Footnote 2).

Research has also clearly demonstrated that OGM is not a simple mood-dependent epiphenomenon of depression. Rather, it appears to represent a stable characteristic of depressed and formerly depressed individuals. So it has been shown that OGM remains stable in spite of recovery from the depression and that it is predictive of an unfavourable course (e.g., Brittlebank, Scott, Williams, & Ferrier, 1993; Peeters, Wessel, Merckelbach, & Boon-Vermeeren, 2002; Raes et al., 2006a). Hence, OGM is regarded as a vulnerability factor for depressive relapse or chronic depression.

Consistent with OGM as a vulnerability factor for depression, it remains present when patients are recovered (e.g., Mackinger, Pachinger, Leibetseder, & Fartacek, 2000a), and predicts emotional (depressed) reactivity to stressful life events in populations that are not suffering from clinical significant pathology. Van Minnen, Wessel, Verhaak, and Smeenk (2005) investigated the relationship between OGM and emotional reactivity to a failed in vitro fertilisation treatment (IVF) in women, measuring memory specificity prior to the stressful event (the bad news). It was found that OGM on the AMT was related to a greater increase in depressive symptoms. A similar pattern emerged in a study by Gibbs and Rude (2004) who found that a higher frequency of negative life events (over a 4–6-week period) led to a significant greater increase in depressive symptoms for students with high levels of OGM at baseline than for low-OGM students, and low-OGM students with low frequencies of stressful life events. Similarly, Mackinger, Loschin, and Leibetseder (2000b) reported that the change in depressive symptomatology from assessment during pregnancy to post-delivery in a non-clinical sample of pregnant women was significantly predicted by memory overgenerality: The more overgeneral memories retrieved during pregnancy, the less reduction or the greater the increase in depressive affect 3 months after delivery (for a replication, see Hipwell, Reynolds, & Crick, 2004). Note that baseline depression levels were controlled for at all times in those prediction studies. Most importantly for the present discussion, it should be noted that the above reviewed studies were all conducted in populations that were not suffering from clinical significant pathology at baseline. Respondents in the Hipwell et al. (2004) study, for example, in which OGM was predictive of depressive symptoms post-delivery, obtained baseline scores on a number of instruments that were comparable with those of ''never-depressed'' non-childbearing controls reported by Teasdale and Cox (2001). In other words, OGM not only constitutes a vulnerability factor for prolonged depression or depressive relapse, but also likely represents a marker of depressed reactivity and, possibly, a vulnerability factor for a first onset of depression in never-depressed individuals.

Given this trait-marker quality of OGM, the AMT might prove a useful tool to screen people at risk for depression or depressive relapse. So in terms of *secondary* prevention, it may help us target preventive interventions to the appropriate groups. As for *primary* prevention, it would then be useful to identify depression-prone people (''overgeneral retrievers'') in *non-clinical groups* as well. However, to date the AMT has rarely been applied in non-clinical populations. The majority of studies that have used the AMT were conducted in clinical, most often depressed or formerly depressed populations. Over the last 5 years, our research group has acquired considerable experience with the use of the AMT in non-clinical groups—in particular in students. Based on the results of a number of our own unpublished studies, there are good reasons for doubting the usefulness of the standard AMT to detect OGM in non-clinical samples.

First of all, the AMT typically leads to a very low frequency of overgeneral memories in non-clinical participants (e.g., students). For example, in three studies that we conducted in students (De Beer, Hermans, & Raes, 2005; Raes & Hermans, 2002, 2003), the mean number of overgeneral (categoric and semantic) memories on a 10-item AMT was only .33 ($SD = .82$; $n = 156$), .21 ($SD = .54$; $n = 97$), and .46 ($SD = .90$; $n = 408$), respectively. One would expect OGM to be more frequent in these groups, given that the disorder for which it is regarded as a vulnerability factor (i.e., depression) is so highly prevalent, especially in university students, who are known to be at relatively high risk for depression (e.g., Rude, Valdez, Odom, & Ebrahimi, 2003). Possibly, the very extensive instructions, the provision of practice trials, and the repeated prompting for specificity in the standard AMT administration jointly mean that people with an overgeneral memory retrieval style in a non-clinical population still manage to obtain high specificity scores

on this task. Or in other words, the task is insufficiently sensitive.

Second, results with non-clinical respondents suggest that memory specificity vs overgenerality as assessed with the AMT is not related to depressive symptomatology. For example, in a group of 97 students, correlations between scores on the Beck Depression Inventory (BDI; Beck, Steer, & Brown, 1996) and the number of specific memories and overgeneral memories retrieved on the AMT were .02 ($p = .85$) and –.07 ($p = .51$), respectively (Raes, Pousset & Hermans, 2004). A median split of scores on the BDI (median = 7) also showed that the low-BDI group did not differ from the high-BDI group with respect to the number of overgeneral memories and the number of specific memories retrieved on the AMT, both ts < 1. The same was true when we compared those scoring below ($n = 72$) and those above ($n = 25$) the widely used cut-off BDI score of 14. If anything, memory specificity is sometimes related to depression in the opposite direction to that expected on the basis of AMT studies in clinical populations, such that those non-clinical individuals performing relatively well on the AMT in terms of memory specificity tend to obtain higher scores on a depression measure. For example, in a study with 156 students (De Beer et al., 2005), correlations between scores on the Center for Epidemiological Studies Depression scale (CES-D; Radloff, 1977) and the number of specific memories and overgeneral memories were .20 ($p < .05$) and –.17 ($p < .05$).

A similar picture seems to be emerging for the relation between depressive rumination and AM specificity. Depressive rumination refers to repeated abstract-conceptual thinking about one's (depressed) feelings and about the possible causes and consequences of these feelings (Nolen-Hoeksema, 1991). Examples of such ruminative thoughts are: ''Why am I feeling depressed?'', ''What could have caused these sad feelings?'', ''Where did it all go wrong?'', ''What if I don't snap out of it?'', etc. Such a ruminative focus on the depression has been convincingly shown to be related to longer and more severe episodes of depression or dysphoria (for a review, see Nolen-Hoeksema, 2004), and thus seems to represent a vulnerability factor for (more severe) depression. Whereas in clinical populations it has been repeatedly found that such a ruminative thinking style is positively related to OGM (Park, Goodyer, & Teasdale, 2004; Raes et al., 2005; Watkins & Teasdale, 2001, 2004; Watkins, Teasdale, & Williams, 2000), in non-clinical populations this relation appears to be absent. For example, in a recent study in 97 students (Raes et al., 2004) respective correlations were .05 ($p = .63$) and –.02 ($p = .92$) between the scores on a rumination measure and the number of specific and overgeneral memories on the AMT, respectively. Likewise, a median split of rumination scores also showed that the low-rumination group did not differ from the high-rumination group with respect to the number of overgeneral memories and the number of specific memories retrieved on the AMT, both ts < 1. The same was true ($t < 1$) when we compared those scoring below ($n = 86$) and those above ($n = 11$) a cut-off rumination score (i.e., the median score on this rumination measure in a sample of patients with major depression). Sometimes the relation between AM specificity and rumination even tends to be opposite to what can be expected based on observations in clinical participants. In a pilot-study (see below), the correlation between a measure of rumination and the number of specific AMT memories was .29 ($p < .09$) in a group of 37 non-clinical respondents.

Why is memory specificity, as measured with the AMT, not related to depression and depressive rumination in non-clinical respondents as it is in depressed ones? As already hinted at above, this may be due to the particular way in which the AMT is being administered (extensive instructions, provision of practice trials, and repeated prompting for specificity). This means that a subgroup of students who have the tendency to retrieve their past in an overgeneral way (and who are thus assumed to be ''at risk'' for depression) still manage to come up with specific memories to cue words when they are explicitly asked to do so and repeatedly instructed *not* to respond with overgeneral material—especially when they are given 1 minute to do so, and when it is explicitly mentioned that their job on this task is to generate specific memories.[3] Once depressed, however, this overgeneral retrieval

[3] And why do these variables, in those cases where they *are* related in non-clinical groups, tend to be related in the opposite direction; that is, the more specific memories respondents retrieve on the AMT, the higher they score on a depression measure and on a rumination scale? Possibly, students who want to do really well on this ''*test*'', e.g., neurotic individuals, will especially try their very best to come up with a specific memory (as requested in this task), and to inhibit all overgeneral thoughts, relative to less neurotic individuals.

style might become a dominant response style, more and more difficult to inhibit. This implies that even when students or other groups of non-clinical participants do respond with specific memories, it may nevertheless be that some of them actually do have a tendency to recall past material in a generalised way; that is, their preferred level of entry into the autobiographical knowledge base, or their habitual level at which they recollect personal information, is situated relatively higher in the hierarchy of autobiographical knowledge, i.e., at the generic-semantic (vs the more specific-episodic) level. The above reviewed data seem to suggest that the AMT, at least in non-clinical groups, is not sufficiently sensitive to detect habitual specificity level of memory recall.[4] Maybe, then, this marker needs to be assessed in non-clinical groups using an alternative method to the AMT. As mentioned, when the AMT is administered, participants are explicitly told that the goal of the task or ''test'' is to generate specific memories. This way, the AMT may miss out the detection of non-clinical respondents' (especially students') habitual or naturalistic style or ''output'' of generative/strategic memory recall (Conway & Pleydell-Pearce, 2000). An alternative procedure to assess the level of overgenerality (vs specificity) with which non-clinical respondents habitually recall personal or autobiographical knowledge would then need to elicit memories without explicitly presenting the task as a memory ''test'', and without instructing and intensively prompting respondents to generate *specific* personal events.

In two pilot studies we explored the usefulness of such an alternative method to elicit or detect an overgeneral retrieval style in non-clinical samples. In a first pilot study ($n = 37$, $M_{age} = 28.9$, $SD = 9.90$), we asked respondents to complete the following sentence stem: ''I still remember well how . . .''. This stem probes people to retrieve past experiences, without the explicit instruction to make it *specific* experiences. In fact, no mention what so ever was made of ''memories'', ''memory test'', or ''recalling the past''. Following the stem completion, participants were administered, in this order, the standard AMT and a rumination scale. A total of 14 participants filled out the incomplete sentence in a specific way (e.g., ''I still remember well how I felt the day I got my driver's licence''), whereas 14 others completed the incomplete sentence with an overgeneral statement (e.g., ''I still remember well how my brother and I used to play with Playmobil® on Sunday mornings''). As predicted, the overgeneral group obtained a higher score on the rumination scale as compared with the specific group, $t(26) = 2.22$, $p < .05$. Thus, this sentence completion procedure led to a considerable amount of overgeneral responses in non-clinical participants, *and* this overgeneral responding was now related to depressive rumination in a similar way to clinical groups. These results were replicated in a second pilot study ($n = 50$; $M_{age} = 21.76$, $SD = 2.25$). Those participants who wrote an overgeneral completion to the sentence stem ($n = 15$) scored significantly higher on a subsequent rumination scale than those who wrote down a specific experience ($n = 35$), $t(48) = 2.07$, $p < .05$.

THE PRESENT STUDIES

Based on the results of these pilot studies, the possibility of such a sentence completion method to assess OGM in non-clinical respondents was further explored in the present studies. In Study 1, a large group of students was administered a series of 11 sentence stems which all probe for past experiences, as well as visual analogue rumination scales, a depression measure, and the standard AMT. Our predictions were (1) that the sentence completion task would elicit more overgeneral responding than the AMT, (2) that OGM as measured with the sentences task would be positively related to depression and depressive rumination, whereas (3) OGM as measured with the standard AMT would not show such a relationship with depression and rumination.

In Study 2, a sentence completion task with the explicit instruction to complete sentences with reference to a specific memory was administered to student participants, alongside the standard completion task from Study 1 (which has no specificity instructions), visual analogue rumination scales, and a depression measure. The aim of Study 2 was to explore whether the enhanced performance of a sentence completion task to assess OGM is due to the fact that it omits the

[4] Absence of any correlation between OGM on the AMT and measures of depression severity might be indicative of the AMT's insensitivity, but this does not necessarily need to be the case. For example, it might be that OGM, as a trait marker for depression, only tends to be present/absent as a function of vulnerability (see our discussion above).

instruction to be specific or because it involves sentence completion. The more detailed hypotheses for Study 2 are summarised at the end of the Discussion section of Study 1.

STUDY 1

Method

Participants

Participants were all first-year psychology students. They participated in return for course credit. As testing was spread out over several time points (see procedural aspects below), only participants who were present during all sessions and who filled out all measures involved in the present study were retained for the analyses. Also, participants who indicated that they had already completed an AMT previously, or that they had not filled out some of the tests seriously (see later), were dropped from the analyses. This resulted in a final group of 197 students (172 women). Their mean age was 18.17 years ($SD = 1.43$; *range* = 17–33).

Materials

Autobiographical Memory Test (AMT). The AMT (Williams & Broadbent, 1986) was used in a written format (see Raes, Hermans, de Decker, Eelen, & Williams, 2003a; Raes, Hermans, Williams, & Eelen, 2006b). Participants are given a booklet with 13 pages. On each page, a cue word is printed. The first two cues are practice items (*grass*, *bread*). The next 10 pages have the test words: *confidence* [*trust*], *scared*, *pleasurable*, *angry*, *courage*, *sad*, *calm* [*at ease*], *bold*, *surprised*, and *stupid*.[5] Participants are given 60 seconds to write down a specific memory for each of the cue words. It is clearly stated in the instructions that a specific memory is one that refers to one particular occasion or event that happened on a particular day at least 1 week before. All cues are embedded in the frame sentence ''Can you write down one specific moment or event that the word X reminds you of?''. When the 60-second time limit for a cue is reached, participants are instructed to turn to the following cue. Cue words and instructions are read aloud by the experimenter.

Each response is coded as a specific or a nonspecific memory. Nonspecific memories are further coded as either an overgeneral categoric memory (e.g., ''going out for a drink''), an overgeneral extended memory (e.g., ''when I worked as a nanny for a year in London''), semantic associate (e.g., ''my brother''), omission, same event (referring to an event already mentioned), or incorrect specific (referring to an event of the past week). Using this scoring procedure, previous studies that were conducted at our lab obtained good reliability (Raes et al., 2003a, 2004) with interrater agreement ranging from 92% to 99% ($K = .83–.96$).

Sentence Completion for Events from the Past Test (SCEPT). The SCEPT comprises 11 sentence stems probing for past experiences. A sample item is ''When I think back to/of . . .''. A list of all items can be found in Appendix 1. Participants are instructed to provide continuations to incomplete sentences (for full instructions, see Appendix 1).

When all sentence stems are completed, participants are instructed to assign a code to each of their responses according to the following sixfold coding system:

> ''1'' (if what you wrote down refers to one specific moment or a particular time); ''M'' (if what you wrote down refers to a repeated activity or a category of similar events without the specification of a particular time [the *M* stands for ''*M*eer dan eens'' or ''*M*eerdere keren'', i.e., ''*M*ore than once'' or ''*M*ultiple times'' in English, respectively]); ''>'' (if what you wrote down refers to an extended period of time which lasted longer than a day); ''?'' (if you seriously tried to complete the sentence, but could not come up with something; i.e., an omission); ''!'' (if you had something in mind, but did not want to write it down for whatever reason, e.g., too personal [besides the exclamation mark, participants are asked in this case to further indicate, using a second code, whether what they had in mind referred to a single event, ''1'', a category of similar events, ''M'', or an extended time period, ''>'']); ''Z'' (if you did not seriously try to complete the sentence in a meaningful way [the *Z* stands for the Dutch ''geen *Z*in om iets serieus te bedenken'',

[5] The Dutch words of the written AMT were *vertrouwen*, *bang*, *prettig*, *boos*, *moed*, *droevig, gerust*, *brutaal*, *verrast*, and *lomp*, respectively. The two Dutch practice items were *gras* and *brood*.

which translates as ''I did not feel like coming up with something serious'']).

Afterwards, sentence completions are coded by the experimenter, using the following coding categories: ''specific memory'' (see ''1'' above); ''categoric memory'' (see ''M'' above), ''extended memory'' (see '' > '' above); and ''omission'' (see ''?'' above). If a respondent wrote down ''!'', then this response is coded as either a specific, a categoric, or extended memory, depending on the second code that this respondent wrote down to further qualify the exclamation mark. A final category is ''semantic associate'' referring to personal overgeneral semantic information (e.g., ''I used to be a very shy girl'', ''I will never forget what my parents mean to me''). (See Appendix 2 for sample responses.) In cases where it is not clear to the experimenter what code a certain response should best be given, then the respondent's own code is taken as the final code. A second independent rater coded the responses of a random selected sample of 20 participants (i.e., 10% of the total sample; 220 sentence completions). Good interrater agreement was shown (87%, $K = .82$).

Beck Depression Inventory (BDI-II). The BDI-II (Beck et al., 1996) is a widely used self-rating measure for severity of depressive symptoms and consists of 21 four-choice statements. Participants are asked to mark the statements that best describe how they felt during the past 2 weeks. The Dutch version by van der Does (2002) is used, for which adequate reliability is reported with Cronbach's alpha of .92 and .93 in a psychiatric and student population, respectively (van der Does, 2002).

Visual Analogue Rumination Scales (VARS). Participants are asked to indicate on a 0 to 10 scale (*not at all* to *very often*) how much they do what is described in four statements *when they are sad, down or feel blue*: ''I have difficulty getting myself to stop thinking about how sad I am''; ''I get absorbed in thinking about why I am sad and find it difficult to think about other things''; ''I repeatedly try to figure out, by doing a lot of thinking, what might be the causes of my sadness''; ''I keep thinking about how I feel, to understand myself and my sad feelings better''. Summing item scores results in a total rumination score (range: 0–40). These four statements were derived from the Leuven Adaptation of the Rumination on Sadness Scale (LARSS; Raes, Hermans, Williams, Bijttebier, & Eelen, in press), which is a revised and extended version of the Rumination on Sadness Scale (RSS; Conway, Csank, Holm, & Blake, 2000). Factor analyses revealed three meaningful factors in the RSS and the LARSS (Raes et al., in press): ''Causal Analysis'' (i.e., ruminating about the possible causes of one's sadness), ''Understanding'' (i.e., ruminating about the meaning of one's sadness), and ''Uncontrollability'' (i.e., thoughts about the uncontrollability of ruminative thinking). Items 1 and 2 of the VARS represent the ''Uncontrollability'' factor of the LARSS. Items 3 and 4 were selected from the ''Causal Analysis'' and ''Understanding'' factor of the LARSS, respectively.

Cronbach's alpha for the four VARS items is .84. Total scores on the VARS correlate highly with scores on widely used rumination scales as the Rumination on Sadness Scale (RSS; Conway et al., 2000) and the Ruminative Response Scale (RRS; Nolen-Hoeksema & Morrow, 1991), respective rs being .82 and .77, both ps $< .001$ (Raes, Hermans, & Eelen, 2003b).

Procedure

All participants were tested collectively and participated in return for course credit. During Session 1, the SCEPT and VARS were administered. Two weeks later, participants completed the BDI-II (Session 2). During Session 3, which took place another 3 weeks later, the AMT was administered. Testing was spread over several sessions to distract participants from guessing the true nature of the study. The SCEPT was administered prior to the AMT to prevent participants from inferring that the sentence completion task is about specificity vs overgenerality of personal past experiences and, thus, to prevent a specific memory response mode (as instructed in the AMT) transfering to the SCEPT.

Results and discussion

Our first hypothesis was that the newly developed sentence completion method (SCEPT) would lead to a higher prevalence of overgeneral responding in this non-clinical sample of students than the standard cue word procedure (AMT).

In support of this hypothesis, the mean proportion of completions that were overgeneral on the SCEPT ($M = .37$, $SD = .15$, range = .09 to .80) was significantly higher than the mean proportion of memories that were overgeneral in response to AMT cue words ($M = .02$, $SD = .04$, range = .00 to .22), $t(196) = 31.18$, $p < .001$. Thus, whereas only a mean of 2% of all responses generated in response to AMT cue words were overgeneral memories, a mean of almost 40% of the sentence completions made on the SCEPT were overgeneral statements.

Our second and third hypothesis proposed, respectively, that OGM as measured with the SCEPT would be positively associated with depression and depressive rumination (Hypothesis 2), and that OGM as measured with the standard AMT would not show such association with depression and rumination (Hypothesis 3). In line with Hypothesis 2, OGM as measured with the SCEPT was significantly correlated with both depression scores (BDI) and depressive rumination scores (VARS), $r(197) = .18$ and $r(197) = .15$, respectively (both ps $< .05$). With respect to Hypothesis 3 the results showed, as predicted, no association between OGM as measured with the AMT and both depression and depressive rumination, respective rs being –.00 ($p = .98$) and –.03 ($p = .57$). The absence of a correlation between OGM on the AMT and both depression and rumination is most likely due to the lack of variance in number of overgeneral memories on the AMT (restriction of range; cf. low base rate of OGM).

All this seems to indicate that, by *not explicitly* asking respondents to recall *specific memories* using a sentence completion procedure, the SCEPT is more sensitive—relative to the AMT—to detect in some non-clinical respondents the tendency to think back on past experiences in an overgeneral way. The fact that OGM on the SCEPT relates to depression and depressive rumination in the same way as in clinically depressed groups is reassuring as to the validity of this method to assess OGM. These findings, then, indicate that OGM as measured with the SCEPT in non-clinical individuals likely reflects the same clinically important cognitive phenomenon as observed in clinical groups where the AMT is used.

However, these results leave unanswered the question of why the SCEPT, as compared to the AMT, produces better data with respect to the detection of OGM in non-clinical respondents. As we have already pointed out in our introduction, we believe that an important reason why the AMT fails to detect OGM in non-clinical respondents is its explicit instruction to recall *specific* memories. As such, one possible and obvious reason why the SCEPT succeeds at eliciting overgeneral responding in some respondents is because it omits the instruction to be specific and encourages some generality (e.g., stems like ''Last year . . .'', ''In the past . . .''). Still, the SCEPT differs from the AMT in yet another respect than the mere presence, respectively absence, of specificity instructions. That is, the SCEPT and AMT also differ in terms of the procedure being used: sentence completion versus free recall to cue words. A second study was carried out to disentangle this confound, and so to evaluate this critical question: Does the SCEPT perform better that the AMT because it omits the instruction to be specific or because it involves sentence completion?

For that purpose, we developed a SCEPT with specificity instruction, which follows the same sentence completion procedure as the standard SCEPT but includes the added explicit instruction for participants to be specific, similar to the standard AMT. In Study 2, respondents filled out this SCEPT with specificity instruction, following the standard SCEPT without specificity instruction (see Study 1), a depression and a rumination measure. In line with the above-mentioned idea that an important reason why the AMT fails to detect OGM in non-clinical respondents is its explicit instruction to recall *specific* memories, we expect that once we add specificity instructions to the SCEPT, this task will lose its ability to detect OGM. This would indeed then show that the SCEPT's improved sensitivity (relative to the AMT) is attributable to the omission of the specificity instruction rather than to its procedural difference (i.e., sentence completion format). More specifically, our predictions thus were (1) that the SCEPT with specificity instruction would elicit less overgeneral responding than the standard SCEPT without specificity instruction, (2) that OGM as measured with the standard SCEPT would be positively associated with depression and rumination, and (3) that OGM as measured with the SCEPT with specificity instruction would show no such associations.

STUDY 2

Method

Participants

A total of 29 first-year psychology students (all women) participated on a voluntary basis. The mean age was 18.40 years ($SD = 0.62$; *range* = 18–20).

Materials

The Sentence Completion for Events from the Past test (SCEPT), the *Beck Depression Inventory (BDI-II), and the Visual Analogue Rumination Scales (VARS) were used as in Study 1.*

Sentence Completion for Events from the Past Test with Specificity Instruction (SCEPT-SI). The same 11 sentence stems were used as in the original SCEPT. However, in the instructions it was now clearly stated that the stems should be completed with reference to a specific memory. The definition of a specific memory is explained in exactly the same way as in the instructions of the AMT. The SCEPT-SI is administered in a written format. For the present study, the SCEPT-SI was administered via the Internet (see Procedure section below for more details).

Procedure

First-year psychology students at the University of Leuven (Belgium) were e-mailed with a request to participate in the study on a voluntary basis. In the e-mail, students were provided with the address of a website where they could take part in the study. Participants were presented the SCEPT-SI, after providing general information (age, sex, and a personalised code that allowed us to link data gathered earlier—SCEPT, BDI-II, and VARS—to the data of this internet survey). A total of 37 students filled out the SCEPT-SI. For the majority (78%) we also had their scores on the earlier administered SCEPT ($n = 29$), and VARS ($n = 29$). Thus, the final sample comprised 29 students. For 23 of these 29 students we also had their scores on the BDI-II. The SCEPT, VARS, and BDI-II were administered, in that order, to first-year psychology students 3 weeks earlier in a group session. They had completed these three questionnaires in return for course credit.

Results and discussion

Our first hypothesis was that the SCEPT without specificity instruction would lead to a higher prevalence of overgeneral responding than the SCEPT with specificity instruction (SCEPT-SI). In line with this hypothesis, the mean proportion of completions that were overgeneral on the SCEPT ($M = .36$, $SD = .13$, range: .18 to .82) was significantly higher than the mean proportion of completions that were overgeneral on the SCEPT-SI ($M = .12$, $SD = .16$, range: .00 to .55), $t(28) = 6.04$, $p < .001$. Thus, whereas for the SCEPT without specificity instruction a mean of 36% of the completions were overgeneral statements, only a mean of 12% of the completions were overgeneral for the SCEPT with specificity instruction.

Our second and third hypotheses contended, respectively, that OGM as measured with the original SCEPT (without specificity instruction) would be positively associated with depression and depressive rumination (Hypothesis 2), and that OGM as measured with the SCEPT with specificity instruction (SCEPT-SI) would show no such association with depression and rumination (Hypothesis 3). In support of Hypothesis 2, OGM as measured with the original SCEPT was significantly correlated with depression scores (BDI-II), $r(23) = .41$, $p < .05$, and marginally significant with depressive rumination scores (VARS), $r(29) = .34$, $p < .07$. Consistent with Hypothesis 3 the results revealed no association between OGM as measured with the SCEPT-SI and both depression and depressive rumination, respective rs being .01 ($p = .97$) and .05 ($p = .79$). Similar to what was the case for the standard AMT in Study 1, levels of OGM using the SCEPT-SI obviously varied insufficiently to detect a relationship with depression and rumination (cf. restriction of range due to low OGM base rate).

The results suggest that, as predicted, the SCEPT is a more sensitive tool for detecting OGM as compared to the AMT because of the omission of the instruction to be specific in the SCEPT rather than because of the sentence completion format (versus free recall to cue words in the AMT).

GENERAL DISCUSSION

The present studies investigated the usefulness of a sentence completion procedure to assess OGM,

a known vulnerability marker for depression. By tradition, researchers have been using a cue word procedure for this purpose, known as the Autobiographical Memory Test (AMT; Williams & Broadbent, 1986). However, whereas the AMT has often been used in clinical populations (see Williams, 2007), it has been applied relatively infrequently in non-clinical populations. Moreover, our own studies showed that the AMT may be, at least in such non-clinical groups, insufficiently sensitive to pick up OGM. For example, the AMT typically leads to a low frequency of overgeneral memories in non-clinical samples. Also, OGM shows a different pattern of correlations in non-clinical than in clinical samples with other variables of relevance such as depression and depressive rumination. As an alternative for the AMT, we developed the Sentence Completion for Events from the Past Test (SCEPT) to measure OGM in non-clinical populations.

In the first study, the SCEPT was administered to a group of first-year university students, alongside a depression scale, a rumination measure, and the traditional AMT. It was found, as predicted, that the sentence completion method led to a significantly higher proportion of overgeneral responses (40%) than the traditional cue word technique (AMT; less than 5%). The results of Study 2, in which additionally a SCEPT was administered that asked participants to complete sentences with *specific* memories, indicated that the SCEPT's superior performance in eliciting OGM in non-clinical participants is most likely due to the omission of the specificity instruction rather than to its procedural difference (i.e., sentence completion format). These results thus suggest that, by not explicitly instructing respondents to retrieve *specific* personal past experiences, the SCEPT is relatively more sensitive to detect non-clinical respondents' tendency to think of past experiences in an overgeneral way.

Importantly, the data of both studies also revealed that level of depression and depressive rumination were positively associated with OGM as measured with the SCEPT, but not when assessed using the AMT (Study 1) or when assessed using a sentence completion with specificity instruction (SCEPT-SI; Study 2). The latter finding is likely due to restriction of range resulting from low base rates of OGM on the AMT and SCEPT-SI (floor effect). The finding that OGM on the SCEPT is related to depression and depressive rumination in non-clinical groups, just as in clinically depressed groups when the AMT is used, further underscores the validity of this method to assess OGM. In other words, OGM as measured with the SCEPT in non-clinical individuals likely reflects the same clinically important cognitive phenomenon as observed in clinical groups where the AMT is used.

The whole of the present results thus suggest that the SCEPT represents a potentially valuable instrument to meaningfully assess OGM in non-clinical individuals. In terms of (primary) prevention, then, the SCEPT may prove useful to identify individuals who are at risk for depression; that is, individuals who have a tendency to recall past events in an overgeneral form. So whereas previously, using the AMT, this known marker for depression or depression-proneness was relatively difficult to detect in non-clinical respondents, the present sentence completion method opens perspectives in this respect.

Finally, a number of limitations of the present studies deserve mention. First, participants in the studies were predominantly women. Thus, it remains to be seen to what extent our findings will generalise to male respondents, although we are not aware of any report of gender differences in the literature on OGM. Second, order of tests was not counterbalanced, neither across testing sessions nor within testing sessions. As such, we cannot exclude possible order effects. However, the fixed order of test presentation, namely that the SCEPT was always presented first, was motivated by the following assumption. If we had administered the standard AMT or SCEPT *with* specificity instructions (SCEPT-SI) before the regular SCEPT (with no specificity instructions), this would undoubtedly have influenced respondents in the way they would respond to the SCEPT. For example, they might assume that they were expected to complete the regular SCEPT stems with *specific* material, similar to the AMT or SCEPT-SI (unintended transference of instructions or response style ''mind set''). This then might cause the SCEPT to miss the detection of one's habitual retrieval style. Third, formal diagnoses (in particular depression and history of depression) were not assessed, leaving unanswered the question whether level of OGM as measured with the SCEPT is higher in clinically depressed or formerly depressed students as compared to students with elevated BDI scores who do not fulfil the full criteria for current or past depression. Future studies, including formal

assessment of clinical diagnosis, are thus warranted.

Future studies should also examine the predictive validity of the SCEPT for emotional distress in response to stressful life events (cf. Gibbs & Rude, 2004), as well as for depression onset and depressive relapse. It will also be important for future research to investigate the validity of the SCEPT in clinical samples. Besides the potential clinical relevance of this sentence completion procedure, the SCEPT may also prove useful to select participants for experimental studies investigating the underlying processes involved in this important cognitive phenomenon. Also, such a sentence completion memory task might be viably used as an evaluation tool for experimental procedures intended to influence OGM. Future studies will show to what extent these ideas apply to real facts.

Manuscript received 2 September 2005
Manuscript accepted 26 February 2007
First published online 31 May 2007

REFERENCES

Barnhofer, T., de Jong-Meyer, R., Kleinpass, A., & Nikesch, S. (2002). Specificity of autobiographical memories in depression: An analysis of retrieval processes in a think-aloud task. *British Journal of Clinical Psychology*, *41*, 411–416.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory – second edition*. San Antonio, TX: The Psychological Corporation.

Brittlebank, A. D., Scott, J., Williams, J. M. G., & Ferrier, I. N. (1993). Autobiographical memory in depression: State or trait marker. *British Journal of Psychiatry*, *162*, 118–121.

Conway, M. A. (1990). *Autobiographical memory: An introduction*. Philadelphia, PA: Milton Keynes.

Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. *Philosophical Transactions of the Royal Society of London, B*, *356*(1413), 1375–1384.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, 594–628.

Conway, M. A., Csank, P. A. R., Holm, S. L., & Blake, C. K. (2000). On assessing individual differences in rumination on sadness. *Journal of Personality Assessment*, *75*, 404–425.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261–288.

Conway, M. A., Singer, J. A., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social Cognition*, *22*, 491–529.

De Beer, E., Hermans, D., & Raes, F. (2005). [*Autobiographical memory specificity and depression in first-year psychology students*]. Unpublished raw data.

Gibbs, B. R., & Rude, S. S. (2004). Overgeneral autobiographical memory as depression vulnerability. *Cognitive Therapy and Research*, *28*, 511–526.

Hipwell, A. E., Reynolds, S., & Crick, E. P. (2004). Cognitive vulnerability to postnatal depressive symptomatology. *Journal of Reproductive and Infant Psychology*, *22*, 211–227.

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, *4*, 268–280.

Mackinger, H. F., Loschin, G. G., & Leibetseder, M. M. (2000a). Prediction of postnatal affective changes by autobiographical memories. *European Psychologist*, *5*, 52–61.

Mackinger, H. F., Pachinger, M. M., Leibetseder, M. M., & Fartacek, R. R. (2000b). Autobiographical memory in women remitted from major depression. *Journal of Abnormal Psychology*, *109*, 331–334.

Nelson, K. D., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, *111*, 486–511.

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, *100*, 569–582.

Nolen-Hoeksema, S. (2004). The response styles theory. In C. Papageorgiou & A. Wells (Eds.), *Depressive rumination: Nature, theory, and treatment* (pp. 107–124). Chichester, UK: John Wiley & Sons.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology*, *61*, 115–121.

Park, R. J., Goodyer, I. M., & Teasdale, J. D. (2004). Effects of induced rumination and distraction on mood and overgeneral autobiographical memory in adolescent major depressive disorder and controls. *Journal of Child Psychology and Psychiatry*, *45*, 996–1006.

Peeters, F., Wessel, I., Merckelbach, H., & Boon-Vermeeren, M. (2002). Autobiographical memory specificity and the course of major depressive disorder. *Comprehensive Psychiatry*, *43*, 344–350.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.

Raes, F., & Hermans, D. (2002). [*Autobiographical memory specificity and depressive rumination in first-year psychology students*]. Unpublished raw data.

Raes, F., & Hermans, D. (2003). [*Autobiographical memory specificity in first-year psychology students*]. Unpublished raw data.

Raes, F., Hermans, D., de Decker, A., Eelen, P., & Williams, J. M. G. (2003a). Autobiographical memory specificity and affect regulation: An experimental approach. *Emotion*, *3*, 201–206.

Raes, F., Hermans, D., & Eelen, P. (2003b). De Nederlandstalige versie van de Ruminative Response Scale (RRS-NL) en de Rumination on

Sadness Scale (RSS-NL) [The Dutch version of the Ruminative Response Scale (RRS-NL) and the Rumination on Sadness Scale (RSS-NL)]. *Gedragstherapie*, *36*, 97–104.

Raes, F., Hermans, D., Williams, J. M. G., Beyers, W., Brunfaut, E., & Eelen, P. (2006a). Reduced autobiographical memory specificity and rumination in predicting the course of depression. *Journal of Abnormal Psychology*, *115*, 699–704.

Raes, F., Hermans, D., Williams, J. M. G., Bijttebier, P., & Eelen, P. (in press). A ''Triple W''-model of rumination on sadness: Why am I feeling sad, what's the meaning of my sadness, and wish I could stop thinking about my sadness (but I can't!). *Cognitive Therapy & Research*.

Raes, F., Hermans, D., Williams, J. M. G., Demyttenaere, K., Sabbe, B., Pieters, G., et al. (2005). Reduced specificity of autobiographical memories: A mediator between rumination and ineffective social problem-solving in major depression? *Journal of Affective Disorders*, *87*, 331–335.

Raes, F., Hermans, D., Williams, J. M. G., & Eelen, P. (2006b). Reduced autobiographical memory specificity and affect regulation. *Cognition & Emotion*, *20*, 402–429.

Raes, F., Pousset, G., & Hermans, D. (2004). *Correlates of autobiographical memory specificity in a nonclinical student population*. Unpublished manuscript.

Rude, S. S., Valdez, C., Odom, S., & Ebrahimi, A. (2003). Negative cognitive biases predict subsequent depression. *Cognitive Therapy and Research*, *27*, 415–429.

Teasdale, J. D., & Cox, S. G. (2001). Dysphoria: Self-devaluative and affective components in recovered depressed patients and never depressed controls. *Psychological Medicine*, *31*, 1311–1316.

Van der Does, A. J. W. (2002). *Handleiding bij de Nederlandse bewerking van de BDI-II* [Manual of the Dutch version of the BDI-II]. San Antonio, TX/ Lisse, the Netherlands: The Psychological Corporation/Swets Test Publishers.

Van Minnen, A., Wessel, I., Verhaak, C., & Smeenk, J. (2005). The relationship between autobiographical memory specificity and depressed mood following a stressful life event: A prospective study. *British Journal of Clinical Psychology*, *44*, 405–415.

Watkins, E., & Teasdale, J. D. (2001). Rumination and overgeneral memory in depression: Effects of self-focus and analytic thinking. *Journal of Abnormal Psychology*, *110*, 353–357.

Watkins, E., & Teasdale, J. D. (2004). Adaptive and maladaptive self-focus in depression. *Journal of Affective Disorders*, *82*, 1–8.

Watkins, E., Teasdale, J. D., & Williams, R. M. (2000). Decentering and distraction reduce overgeneral autobiographical memory in depression. *Psychological Medicine*, *30*, 911–920.

Williams, J. M. G. (2000). *Autobiographical Memory Test: Report of the AMT consensus meeting*. Institute of Medical and Social Care Research, University of Wales, Bangor, UK.

Williams, J. M. G., Barnhofer, T., Crane, C., Hermans, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin*, *133*, 122–148.

Williams, J. M. G., & Broadbent, K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*, *95*, 144–149.

Williams, J. M. G., & Dritschel, B. H. (1992). Categoric and extended autobiographical memories. In M. A. Conway, D. C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspectives on autobiographical memory* (pp. 391–412). Dordrecht, the Netherlands: Kluwer Academic.

APPENDIX 1

Dutch and English instructions and items of the Sentence Completion for Events from the Past Test (SCEPT).

Instructies/*Instructions*

Zinaanvultaak
Hieronder vindt u elf zinnen. Eigenlijk zijn het maar halve zinnen. Want wat er telkens staat geschreven, is enkel het begin van een zin. Het is de bedoeling dat u elke zin verder aanvult. U mag de zinnen aanvullen zoals u wil, zolang het maar aansluit bij hetgeen er al geschreven staat. Let er ook op dat u zorgt dat elke zin over iets anders gaat.

Sentence completion task
Below you will find eleven sentences. Actually these are only parts of sentences, because only the beginning of each of the sentences is provided. The purpose of the task is for you to complete each of the sentences. You can complete the sentences any way you want, just as long as what you write corresponds to the provided stem. Also make sure that each of the sentences is on a different topic.

Items/Items

1. Ik herinner me nog goed hoe . . ./*I still remember well how . . .*
2. Ik weet nog dat ik . . ./*I still recall how/that I . . .*
3. Vorig jaar . . ./*Last year . . .*
4. Vroeger . . ./*In the past . . .*
5. Vorige week heb ik . . ./*Last week I . . .*
6. Ik zie nog zo voor me hoe . . ./*I can still picture how . . .*
7. Als ik terugdenk aan . . ./*When I think back to/of . . .*
8. Ik zal nooit vergeten . . ./*I will never forget . . .*
9. Het belangrijkste dat ik ooit heb . . ./*The most important thing that I have ever . . .*
10. Vorig jaar heb ik . . ./*Last year I . . .*
11. Toen ik . . ./At the time when I . . .

APPENDIX 2

Sample responses for different coding categories for the Sentence Completion for Events from the Past Test (SCEPT).

''Specific memory''

I still remember well how . . . sad I was the day my grandfather died.
I will never forget . . . that a friend threw me a surprise party when I turned sixteen.
Last week . . . I held my baby nephew in my arms for the very first time.

''Categoric memory''

I can still picture how . . . my grandmother used to play games with me when I was little.
Last year . . . I went to school by bike everyday.
In the past . . . I used to avoid other people at social gatherings.

''Extended memory''

When I think back to/of . . . my time in junior high, I feel happy.
Last year I . . . went on scout camp for a week as a cook.
I still recall that I . . . was ill for two weeks in a row last year.

''Semantic associate''

In the past . . . I was a very shy person.
The most important thing that I have ever . . . had and have, is my family.
In the past . . . I had short hair.