The Contribution of Metamemory Deficits to Schizophrenia

Steffen Moritz
University of Hamburg

Todd S. Woodward
Riverview Hospital and Simon Fraser University

A number of recent studies have demonstrated that individuals with schizophrenia display knowledge corruption; that is, they hold false information with strong conviction. This aberration in metamemory is thought to stem from poor memory accuracy in conjunction with impaired discrimination of correct and incorrect judgments in terms of confidence. Thirty-one participants with schizophrenia, along with 61 healthy control participants and 48 control participants with other psychiatric conditions, participated in a computerized source memory task. Whereas no differences in memory accuracy were observed between the group with schizophrenia and the group with other psychiatric diagnoses, knowledge corruption was specifically impaired in those with schizophrenia. Schizophrenia participants showed a significantly decreased confidence gap: They were more confident in errors and less confident in correct responses relative to those in the control groups. Knowledge corruption is theorized to be a potential risk factor for the emergence of delusions.

Keywords: schizophrenia, delusions, metamemory, source monitoring, source memory

By international consensus, an episode of schizophrenia is characterized by a severe distortion of a person’s sense of reality. Although a diagnosis of schizophrenia can be based on the presence of disorganized and negative symptoms alone, delusions are considered the core of the disorder and are given special weight in current diagnostic criteria. In sharp contrast to early pathogenetic models contending that a delusion proper must defy understanding (Jaspers, 1973; Walker, 1991), several cognitive mechanisms have recently been targeted that may explain the emergence of fixed false beliefs (for reviews, see Blackwood, Howard, Bentall, & Murray, 2001; Garety & Freeman, 1999). These putative underpinnings of the disorder appear to operate on neutral scenarios that are not (yet) contaminated by a paranoid belief system. For example, using neutral material, it has been repeatedly demonstrated that individuals with schizophrenia—presently deluded or nondeluded—tend to use little information when making firm judgments (Dudley, John, Young, & Over, 1997; Garety, Hemsley, & Wessely, 1991; Huq, Garety, & Hemsley, 1988; Moritz & Woodward, 2005) and demonstrate a bias against disconfirmatory evidence (Woodward, Moritz, Cuttler, & Whitman, in press). Similarly, theory-of-mind deficits are ubiquitous in schizophrenia (Brune, 2005; Frith & Corcoran, 1996), although the specificity of this dysfunction to schizophrenic delusions is subject to ongoing debate (Greig, Bryson, & Bell, 2004; Sarfati, Hardy-Bayle, Besche, & Widlocher, 1997).

Memory and Metamemory in Schizophrenia

In recent years, evidence has mounted that individuals with schizophrenia display excessive knowledge corruption; that is, a high proportion of what patients strongly believe to be true is in fact false (Moritz & Woodward, 2002; Moritz, Woodward, Cuttler, Whitman, & Watson, 2004; Moritz, Woodward, & Ruff, 2003; Moritz, Woodward, Whitman, & Cuttler, 2005). Knowledge, according to this view, refers to memories or pseudomemories that an individual holds with strong conviction (Hemsley & Garety, 1986, p. 52). Correspondingly, a memory episode—irrespective of its validity—in which an individual is not confident is not considered knowledge. Although healthy participants are also fallible (i.e., commit errors) and mistake self-construed mental images as memories (Loftus, 2003; Roediger & McDermott, 1995), memory errors are usually accompanied by a feeling of doubt (i.e., not trustworthy tags are applied to these memory representations), which prevents such cognitions from becoming strong cues for behavior and social interaction. When recalling memory episodes, healthy participants are attempting to retrieve item-specific information (Roediger, Watson, McDermott, & Gallo, 2001) and base their judgments on the vividness of the memory, which in many cases represents a valid heuristic for differentiating an episode as valid or invalid (Reisberg, 2001, p. 211).

Knowledge corruption in schizophrenia is the product of two major sources: an increased number of memory errors in combination with difficulties in discriminating correct and incorrect responses in terms of confidence. We refer to the latter as a reduced confidence gap (Moritz et al., 2003): Those with schizophrenia are both more confident in incorrect responses and less confident in correct responses compared with participants in control groups. In addition to our studies, overconfidence in errors for schizophrenia patients has also been noted in semantic memory tasks (Danion, Gokalsing, Robert, Massin-Krauss, & Bacon, 2001).
Memory is the cognitive function most severely affected by schizophrenia (Heinrichs & Zakzanis, 1998), and memory impairments have long been suspected to contribute to the emergence of the schizophrenic illness (Hemsley, 1994). However, a mere memory-based pathogenetic account of schizophrenia is not fully conclusive, because individuals with dementia or depression, who are also compromised on memory performance (Duffy & O’Carroll, 1994; Moritz, Heeren, Andresen, & Krausz, 2001), do not typically display psychotic symptoms. Moreover, the functional relevance of memory deficits for the establishment of paranoid ideation has remained vague. The synthesis of metamemory (e.g., memory confidence, remember–know judgments) and objective memory performance may prove useful to overcome these empirical and theoretical shortcomings.

In the following, we stress our argument that memory errors and a reduced confidence gap are important contributors to knowledge corruption. However, it is important to note that aberration in one but not the other may be insufficient to generate knowledge corruption. To demonstrate, consider a healthy person with excellent memory abilities who is aware that his or her memory is very reliable. This person may be overly confident in his or her infrequent memory errors. This would, however, not lead to excessive knowledge corruption, because a high proportion of the knowledge system would still consist of accurate memories. Likewise, an individual with dementia who commits many memory errors, but who is aware of the fallibility of his or her memory, would also not suffer from severe knowledge corruption because of overall low memory confidence.

It has been demonstrated in some studies that those with obsessive–compulsive disorder (OCD) or depression are somewhat underconfident (Zitterl et al., 2001) about their individual level of memory functioning and even seem to exaggerate their deficits (Moritz, Ferahli, & Naber, 2004). Those with Alzheimer’s disease, on the other hand, are often aware of their severe cognitive decline and therefore show correspondingly little confidence in their memory. However, those with schizophrenia, while having deficits on memory accuracy comparable with those of depressed individuals and having even better memory in relation to those with Alzheimer’s disease, are less cognizant of their deficits (Bacon, Danion, Kauffmann-Muller, & Bruant, 2001, p. 474; Moritz, Ferahli, & Naber, 2004) and are therefore less cautious in judging the integrity of their perceptions and memories.

In a recent study (Moritz, Woodward, et al., 2005), we demonstrated that while knowledge corruption in schizophrenia is particularly pronounced for false-negative errors (i.e., presented information misjudged as new), knowledge corruption is also detectable for false-positive errors (i.e., new information misjudged as previously presented). In another study (Moritz, Woodward, et al., 2004), excessive knowledge corruption was confirmed for false-negative errors. In the same study, those with schizophrenia also showed a trend toward greater knowledge corruption in false-positive errors when baseline memory performance was taken into account.

The Present Study

In the present study, we pursued three major aims. First, we wanted to establish the specificity of knowledge corruption to schizophrenia. Prior studies compared those with schizophrenia exclusively with healthy controls, leaving open the possibility that knowledge corruption is a feature attached to psychiatric illness, rather than one confined to schizophrenia. In the present study, participants with OCD and posttraumatic stress disorder (PTSD) served as controls. These two disorders are particularly interesting control groups for a study that aimed to elucidate a potential mechanism of reality distortion. This is because, on the one hand, individuals with OCD and PTSD display symptoms that are reminiscent of positive schizophrenia symptomatology (e.g., intrusions of unwanted thoughts and images that severely interfere with daily activities), but on the other hand, they show metacognitive awareness that these thoughts and images arise from their own minds. It was hypothesized that knowledge corruption is significantly stronger in individuals with schizophrenia relative to healthy individuals and those with other psychiatric conditions.

Second, along with confidence judgments, participants were asked to carry out remember–know judgments. This measure was administered because Danion and coworkers (Danion, Kazes, Huron, & Karchouni, 2003; Danion, Rizzo, & Bruant, 1999; Huron & Danion, 2002) previously demonstrated that individuals with schizophrenia make fewer remember ratings compared with controls, suggesting that their memories are less vivid. The joint assessment of these two metamemory dimensions (confidence ratings and remember–know judgment) allowed us to test whether participants with schizophrenia were more inclined to rely on familiarity than on conscious recollection (Moritz, Woodward, et al., 2004; Weiss, Dodson, Goff, Schacter, & Heckers, 2002) and therefore made high-confidence judgments on the basis of suboptimal and/or inaccurate evidence. Whereas it was hypothesized that remember ratings would be associated with higher confidence ratings regardless of group membership, we hypothesized higher confidence ratings for know ratings in individuals with schizophrenia relative to healthy individuals and those with other psychiatric conditions. Such a finding would lend further support to Garety et al.’s (1991) jumping-to-conclusions account, claiming that those with schizophrenia have a data-gathering bias toward quickly accepting hypotheses that are not sufficiently supported for such strong endorsement (Moritz & Woodward, 2004).

Third, in source-monitoring studies, an association between hallucinations and a tendency to misremember an internally generated event as external has been repeatedly reported for participants with schizophrenia (Bentall, Baker, & Havers, 1991; Brebion et al., 2000; Woodward, Menon, & Hu, 2005) and for healthy participants who score high on measures tapping hallucinatory experiences (Larø, Van der Linden, & Marczewski, 2004). However, failures to replicate this relationship outnumber these findings (Harvey, 1985; Keefe, Arnold, Bayen, McEvoy, & Wilson, 2002; Keefe, Poe, McEvoy, & Vaughan, 2003; Moritz & Woodward, 2002; Moritz et al., 2003; Morrison & Haddock, 1997; Seal, Crowe, & Cheung, 1997; Vinogradov et al., 1997). To account for the ambiguity in the literature, it has been proposed that such a bias may operate only with negatively valenced material (Morrison & Haddock, 1997) and when cognitive load is high (Larø et al., 2004). Specifically, externalization of one’s thoughts may occur only for cognitions that are perceived as harmful, taboo, or inappropriate by participants. To provide a fair test of this revised account, the present study, unlike our forerunner investigations, used a source memory test presenting negatively valenced, positively valenced, and neutral material. Moreover, we used cued
anagrams to enhance cognitive load for self-generated items. We hypothesized that participants with schizophrenia who experienced hallucinations would show a bias to misattribute self-generated negatively valenced information to the external source.

Method

Participants

Thirty-one inpatients with an established diagnosis of schizopreniform disorder or schizophrenia according to criteria from the Diagnostic and Statistical Manual of Mental Disorders (4th ed. [DSM-IV]; American Psychiatric Association, 1994) took part in the experiment. For the 2 patients with schizopreniform disorder, a schizophrenia spectrum diagnosis was confirmed at discharge. One control group included 28 patients with PTSD and 20 patients with OCD. All patients were screened by trained clinicians for validity of suspected diagnosis prior to further investigation. Preselected patients then underwent a thorough psychopathological assessment using the MINI International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). The schizoprenic and PTSD psychopathology was further evaluated with the corresponding sections of the German version of the Structured Clinical Interview for DSM-IV (SCID; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997). Through advertisement, an established participant pool, and word of mouth, 61 healthy individuals were recruited for participation in another control group and were carefully screened with the MINI interview for the absence of any psychiatric disorder, including drug and alcohol dependence.

None of the participants with schizophrenia had a known history of brain damage, severe substance abuse, or an Axis I diagnosis other than schizopreniform disorder or schizophrenia, except for depression, which was tolerated. Participants in the control groups did not display any current or past psychotic symptoms (i.e., delusions, hallucinations). Without knowledge of the neurocognitive results, we excluded OCD patients with a lack of insight or overvalued ideas, as assessed with the Yale–Brown Obsessive–Compulsive Scale (Y-BOCS; Goodman et al., 1989; Hand & Büttner-Westphal, 1991; Jacobsen, Kloss, Fricke, Hand, & Moritz, 2003). This led to the exclusion of 5 participants prior to assessment.

Schizoprenic psychopathology was assessed with the Positive and Negative Syndrome Scale (PANSS; Kay, Opler, & Lindenmayer, 1989). The PANSS was complemented with six additional ratings derived from the Positive and Negative and Disorganized Symptoms Scale (PANADSS; Moritz, Andresen, et al., 2001) tapping inadequate affect, flat affect (not separated in the PANSS), associative loosening, thought blocking (not separated in the PANSS), auditory hallucinations, and other hallucinations (the PANSS does not distinguish between different types of hallucinations). Because the ratings were based on the original semistructured interviews of the scales, the excellent interrater reliabilities of the PANSS and PANADSS (Gottsch, 2005) are applicable to the ratings used in the present investigation. However, interrater reliability was not independently assessed.

Three syndromes were composed from the PANSS/PANADSS, which closely followed recent factor analytic solutions obtained with these scales (Mass, Schoemig, & Hiltschfeld, 2000; Moritz, Andresen, et al., 2001): the positive syndrome (delusions, hallucinations, suspiciousness/ideas of persecution, unusual thought content), the negative syndrome (flat affect, emotional withdrawal, lack of relationship, passive social withdrawal, lack of spontaneity), and disorganization (associative loosening, inadequate affect, problems with abstract thinking, attention, disorientation). The majority of participants with schizophrenia displayed at least mild delusional symptoms at the time of testing (n = 25, 81%). Hallucinations were reported by 14 participants (45%), and approximately half of the participants with schizophrenia had at least mild symptoms of formal thought disorder (n = 15, 48%). The severity of the OCD symptomatology in the 20 participants with OCD was assessed with the Y-BOCS (M = 22.55, SD = 5.68). Of the participants with OCD or PTSD, 24 (50%) fulfilled criteria for a current episode of major depression or dysthymia. Additional psychopathological characteristics are displayed in Table 1. After complete description of the study to the participants, written informed consent was obtained.

This article partly overlaps with another of Steffen Moritz’s investigations studying source memory biases in OCD. The reporting of these studies is not merged, because they have distinct research goals and draw on a different body of literature.

Procedure

Subsequent to the psychopathological assessment, a computerized source-monitoring task was administered using Superlab (Cedrus Corporation, 1990) for the Macintosh. Participants were seated in front of a computer monitor and were told that they would be presented with word puzzles (i.e., anagrams). The letters of the to-be-identified words (written in red uppercase letters) were jumbled by randomly selecting one of the letters, and displacing it to any other location in the word (e.g., tebla; solution: table). Above each word puzzle, a short written verbal clue in

Table 1

Sociodemographic and Psychopathological Characteristics of the Samples

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy (n = 61)</th>
<th>Schizophrenia (n = 31)</th>
<th>OCD or PTSD (n = 48)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>31.05</td>
<td>8.75</td>
<td>33.77</td>
<td>9.90</td>
</tr>
<tr>
<td>Years of school</td>
<td>11.95</td>
<td>1.52</td>
<td>11.61</td>
<td>1.65</td>
</tr>
<tr>
<td>Psychopathological variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of hospitalizations (including present)</td>
<td>68.45</td>
<td>17.48</td>
<td>1.29</td>
<td>1.22</td>
</tr>
<tr>
<td>PANSS total score</td>
<td>4.81</td>
<td>4.82</td>
<td>1.29</td>
<td>1.22</td>
</tr>
<tr>
<td>HDRS total score</td>
<td>0.87</td>
<td>2.70</td>
<td>14.72</td>
<td>7.13</td>
</tr>
<tr>
<td>Chlorpromazine equivalent dosage in mg</td>
<td>687.88</td>
<td>864.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. For the healthy group, 31 were men and 30 were women; for the group with schizophrenia, 22 were men and 9 were women; for the group with obsessive-compulsive disorder (OCD) or posttraumatic stress disorder (PTSD), 27 were men and 21 were women; χ²(1, N = 140) = 3.43, p > .1. Nine participants in the group with schizophrenia and 25 participants in the group with OCD or PTSD were taking antidepressant agents; χ²(1, N = 48) = 6.13, p = .01. PANSS = Positive and Negative Syndrome Scale; HDRS = Hamilton Depression Rating Scale.

* All but 3 of the participants with schizophrenia received neuroleptic treatment.
green font was provided to disambiguate the puzzles (e.g., furniture). The methodology for this task was adopted from a pilot project (Woodward, Whitman, Cuttler, & Moritz, 2003). Participants were familiarized with the procedure that, in alternating order, the word puzzles should be solved either by the participant or the computer, which was indicated by a brief computer display (Please name the next item or The next item will be read to you). In the case of computer-generated trials, a prerecorded female voice was played, providing the correct answer immediately prior to the word puzzle, ensuring that the word puzzle could not be solved by the participant in advance. Participants were instructed not to verbalize the computer-generated answer, but to pay attention to the clue, the word puzzle, and the correct answer. In total, 48 word puzzles were presented: 24 were solved by the participant, and 24 were solved by the computer. Each set of 24 items was evenly divided across three affective categories, that is, 8 positive (e.g., feast, talent), 8 neutral (e.g., hat, apple), and 8 negative (e.g., war, loss).

In direct succession to the learning trial, participants were presented with a recognition trial. The recognition trial consisted of 96 visual items: 48 old items (i.e., solutions to the word puzzles) and 48 new items, presented in random order. Recognition items were written in lowercase blue font to prevent physical matching with encoding items. Half of the 48 new items were semantically related to the old items (i.e., of the 24 new items, 12 were related to self-generated items, and 12 were related to computer-generated items). Related distractor items were used to test the memory more demanding, because the computation of some indices relied on the occurrence of errors for each participant. The 48 distractor items were again evenly distributed over the 3 affective categories (i.e., 16 positive, 16 negative, and 16 neutral words). The stimulus material was derived from a pool of 120 words that were rated for valence on a 7-point Likert scale (1 = very negative, 4 = neutral, 7 = very positive) by 20 raters. Words were regarded as negative if their mean rating was below 3, positive if their mean rating was above 5, and neutral if their mean rating fell between 3.5 and 4.5. Words from the affective categories were equated on emotional valence, in that they were rated the same distance from the mean (p > .1).

The experimental procedure is displayed in Table 2.

For the recognition task, participants were asked to provide three answers in response to each recognition item. First, participants specified whether the item was computer generated, self-generated, or new. Second, participants were required to evaluate, on a 4-point Likert scale, the degree to which the item was computer generated, self-generated, or new. The main metamemory parameters of interest are discussed below.

Table 2
Number of Trials Separated by Word Type for the Learning (48 Items) and Recognition Phase (96 Items)

<table>
<thead>
<tr>
<th>Valence</th>
<th>Old (learning)</th>
<th>New (recognition)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Computer</td>
<td>Self</td>
</tr>
<tr>
<td>Positive</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Negative</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Neutral</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Data Analysis

The main metamemory parameters of interest are discussed below.

Confidence in responses. Mean confidence in errors was subtracted from mean confidence in correct responses to gain an index of the confidence gap, that is, the degree to which a participant was able to discriminate true from false memories in terms of confidence. A low score thereby indicates that a participant was similarly confident about responses that turned out either to be correct or incorrect. In healthy participants, correct responses are usually associated with higher confidence, whereas incorrect responses are cognitively marked with not trustworthy tags, leading to lower confidence ratings. The confidence-gap measure was used because in past studies, we observed that this measure was more sensitive to group differences than the individual mean differences on incorrect and correct responses between participants with schizophrenia and healthy participants (Moritz et al., 2003; Moritz, Woodward, et al., 2005). Consistent with past work (Moritz & Woodward, 2002; Moritz, Woodward, et al., 2004, 2005), we also computed an alternative confidence-gap parameter based on the percentage of responses made with high confidence (i.e., a confidence rating of 4) for correct and incorrect responses.

Knowledge corruption index. The percentage of high-confidence responses that were errors was calculated (the knowledge corruption index). This index measures the degree to which false information intrudes into the knowledge system (i.e., what one believes to be factual). In past studies, when no high-confidence judgments were made at all, the knowledge corruption index was set to zero (despite a zero in the denominator), because knowledge cannot be regarded as corrupted when no high-confidence judgments are made. However, in the current study, this did not occur for any of the participants.

Results

Sociodemographic and Background Variables

The three samples did not differ with respect to gender distribution, age, or years of education (see Table 1). However, more psychiatric hospitalizations were recorded for the participants with schizophrenia compared with the control group with OCD or PTSD. Moreover, there were (nonsignificantly) more male participants in the schizophrenia group. For this reason, number of hospitalizations and gender were carefully considered as potential moderator variables, but these variables did not modify any of the cognitive effects reported below.

Memory Performance (Accuracy)

A two-way analysis of variance (ANOVA) was conducted with group (schizophrenia, healthy, OCD/PTSD) as the between-subjects variable and type of memory error (false positive, false negative, source memory error) as the within-subjects variable. Number of errors served as the dependent variable. The main effects of group, $F(2, 137) = 5.35, p = .006$, and type of memory
error, $F(2, 137) = 26.22, p < .001$, achieved significance as did the Group $\times$ Type of Memory Error interaction, $F(4, 274) = 3.20, p = .03$. The significant main effect for group was due to an overall higher error rate in the groups with schizophrenia or OCD/PTSD relative to healthy participants (see Table 3). The significant effect of type of memory error was due to overall higher false-negative errors (i.e., forgetting) relative to false-positive and source memory errors (both $ps < .001$), whereas the numbers of memory errors for the latter types were not significantly different ($p > .3$). The origin of the significant interaction is evident from Table 3: The groups with schizophrenia or OCD/PTSD made more false-negative errors than did healthy participants (post hoc $t$ tests were $p < .01$ or lower). No group differences emerged on the overall number of source memory errors and false-positive errors. When valence was entered in the above ANOVA as an additional within-subject variable, the results remained essentially unchanged ($p > .1$). Specifically, there was no evidence that the tendency to externalize negative self-generated information differed among the groups, $F(2, 137) = 0.06, p > .9$.

A signal detection analysis (Macmillan & Creelman, 1990) based on accuracy scores and the number of false positives (see the Discussion section) revealed that the groups with schizophrenia or OCD/PTSD displayed an attenuated sensitivity to the distinction between new and old items, indicated by decreased $d'$ scores (at least $p = .05$) relative to healthy controls (healthy participants: $M = 2.03$; participants with schizophrenia: $M = 1.67$; participants with OCD or PTSD: $M = 1.80$). In addition, the participants with schizophrenia ($M = 0.37$) showed a larger beta score than healthy participants ($M = 0.13, p = .01$), but did not differ from the participants with OCD or PTSD ($M = 0.23$), a strategy possibly adopted to avoid the tendency to commit false positives when sensitivity is reduced.

### Confidence Gap

To determine whether patients with schizophrenia were less able to discriminate correct from incorrect memory responses in terms of confidence, an ANOVA was conducted with group (schizophrenia, healthy, OCD/PTSD) as the between-subjects variable and response type (correct, incorrect) as the within-subjects variable. Mean confidence in responses served as the dependent variable. The group effect did not achieve significance, suggesting comparable overall memory confidence ($p > .7$). Results did not change when the group with OCD and PTSD was split: OCD patients did not show decreased confidence ratings ($p > .5$). The effect for response type was highly significant, indicating that participants were more cautious overall (i.e., gave lower confidence ratings) when making incorrect responses, $F(1, 137) = 294.45, p < .001$. In agreement with prior studies (Moritz & Woodward, 2002; Moritz, Woodward, et al., 2005), the Response Type $\times$ Group interaction was significant, $F(2, 137) = 9.91, p < .001$, indicating a between-groups difference in the confidence gap. As is displayed in Figure 1, participants with schizophrenia displayed the smallest confidence gap: They were less able to discriminate correct and incorrect responses in terms of confidence. Post hoc analyses demonstrated that the effect was significant compared with both healthy participants and those with OCD or PTSD ($p < .001$). When the group with OCD or PTSD was divided into its two subsamples, participants with schizophrenia displayed a significantly attenuated confidence gap compared with participants with PTSD ($p = .01$) and OCD ($p < .001$). The decreased confidence gap was due to two influences: Relative to healthy participants and those with OCD or PTSD, participants with schizophrenia displayed underconfidence for correct responses ($M$s = 3.23, 3.18, 3.13, respectively) while being at the same time overconfident in

### Table 3

<table>
<thead>
<tr>
<th>Source memory task</th>
<th>Healthy (H; $n = 61$)</th>
<th>Schizophrenia (S; $n = 31$)</th>
<th>OCD or PTSD (O/P; $n = 48$)</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total errors</td>
<td>24.51, 8.42</td>
<td>30.87, 10.11</td>
<td>28.06, 9.39</td>
<td>$F(2, 137) = 5.35, p = .006; S, O/P &gt; H (p = .002, $p = .05$)</td>
</tr>
<tr>
<td>Source attribution errors</td>
<td>7.67, 4.17</td>
<td>8.45, 4.65</td>
<td>7.60, 4.06</td>
<td>$F(2, 137) = 0.44, p &gt; .6, ns$</td>
</tr>
<tr>
<td>Remember ratings for source attribution</td>
<td>1.67, 1.92</td>
<td>3.00, 3.94</td>
<td>1.58, 1.89</td>
<td>$F(2, 137) = 3.63, p = .02; S &gt; H, O/P (p = .02, $p = .01$)</td>
</tr>
<tr>
<td>False-positive errors</td>
<td>7.49, 5.81</td>
<td>7.32, 6.65</td>
<td>7.40, 4.95</td>
<td>$F(2, 137) = 0.10, p &gt; .9, ns$</td>
</tr>
<tr>
<td>Remember ratings for false-positive errors</td>
<td>0.82, 1.89</td>
<td>1.29, 2.08</td>
<td>0.69, 1.52</td>
<td>$F(2, 137) = 1.10, p &gt; .3; ns$</td>
</tr>
<tr>
<td>False-negative errors</td>
<td>9.34, 4.76</td>
<td>15.10, 8.29</td>
<td>13.06, 9.35</td>
<td>$F(2, 137) = 7.09, p = .001; S, O/P &gt; H (p = .001, $p = .01$)</td>
</tr>
<tr>
<td>Knowledge corruption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total errors (%)</td>
<td>12.62, 8.50</td>
<td>21.90, 12.06</td>
<td>14.84, 11.73</td>
<td>$F(2, 137) = 8.10, p &lt; .001; S &gt; O/P, H (p &lt; .005, $p &lt; .001$)</td>
</tr>
<tr>
<td>Source attribution errors (%)</td>
<td>4.47, 5.00</td>
<td>6.94, 7.86</td>
<td>3.88, 4.11</td>
<td>$F(2, 137) = 3.13, p &lt; .05; S &gt; H, O/P (both $p &lt; .05$)</td>
</tr>
<tr>
<td>False-positive errors (%)</td>
<td>3.28, 3.80</td>
<td>3.40, 4.17</td>
<td>2.69, 3.67</td>
<td>$F(2, 137) = 0.43, p &gt; .6, ns$</td>
</tr>
<tr>
<td>False-negative errors (%)</td>
<td>4.87, 4.85</td>
<td>11.55, 12.19</td>
<td>8.27, 11.28</td>
<td>$F(2, 137) = 5.53, p = .005; S &gt; H (p = .001$)</td>
</tr>
</tbody>
</table>

Note. OCD = obsessive–compulsive disorder; PTSD = posttraumatic stress disorder.
errors (M = 2.69, 2.68, 2.86, respectively). However, the contrast between participants with schizophrenia and the control groups achieved significance (p ≤ .05) only for incorrect responses.

A related analysis using the percentage of high-confidence judgments as the dependent variable yielded similar results (correct responses: M = 49% for healthy controls, 50% for psychiatric controls, and 47% for those with schizophrenia, respectively; incorrect responses: M = 21%, 24%, 29%, respectively). As was the case for the analysis involving mean confidence, the Group × Response Type interaction was highly significant, F(1, 137) = 5.38, p = .006, and post hoc analyses confirmed that the distinctiveness of correct and incorrect responses was significantly reduced for participants with schizophrenia compared with both healthy participants and those with OCD or PTSD (p < .01).

Knowledge Corruption

As can be seen in Table 3, participants with schizophrenia showed increased total knowledge corruption relative to both control groups. The significant difference was maintained when participants with OCD and PTSD were compared separately to schizophrenia patients (OCD: p = .01; PTSD: p = .02). On closer inspection, knowledge corruption was increased for participants with schizophrenia on source confusions (relative to both control groups) and for false-negative judgments (relative to healthy controls). No significant differences emerged with respect to false-positive errors (i.e., misjudging new information as either self- or experimenter generated).

A subsidiary analysis revealed that the group difference for the knowledge corruption index on source attribution errors was mainly attributable to computer-generated items being misperceived as self-generated: schizophrenia, 4.49% (SD = 6.36); healthy, 1.98% (SD = 2.88); OCD or PTSD, 2.05% (SD = 3.49); F(2, 137) = 4.43, p = .01; for all post hoc analyses, p < .01. The knowledge corruption index for false externalizations (i.e., attribution to computer instead of self) did not differ between groups, F(2, 137) = 0.38; p > .6, as was the case for the knowledge corruption index for externalizing negative self-generated information, F(2, 137) = 1.65, p > .1.

Figure 1. Means and standard deviations of the confidence gap (mean confidence correct minus mean confidence incorrect). Participants with schizophrenia were the least able to discriminate correct from incorrect responses in terms of confidence. That is to say, compared with both control groups, participants with schizophrenia demonstrated a combination of underconfidence in correct responses and overconfidence in errors. ***p ≤ .001.

Remember–Know Judgments

Table 3 shows the data for the frequency of incorrect remember judgments. Overall, participants with schizophrenia made more incorrect remember judgments (difference from healthy controls: p = .02; difference from psychiatric controls: p = .03), which was primarily due to a greater number of errors for source misjudgments (see Table 3).

In Table 4, the relative distribution of remember, know, and new ratings is displayed across the different conditions. A three-way mixed ANOVA was conducted with group (schizophrenia, healthy, and OCD or PTSD) as the between-subjects variable and item type (source item, new-related item, new-unrelated items) and response type (remember, know) as within-subjects variables. Percentages of response ratings classified as remember or know judgments served as the dependent variables. To conduct one comprehensive analysis, we collapsed the source judgments that were correct and those that were incorrect into one category to avoid unbalanced means caused by different numbers of correct and incorrect responses (see Table 4). As expected, a significant effect of item type emerged, F(2, 274) = 1,373.78, p < .001. Follow-up contrasts showed that participants overall made the fewest remember and know ratings for the unrelated condition, followed by the related condition and the source items (at least p < .05). This effect was further moderated by response type as evidenced by a significant Item Type × Response Type interaction, F(1, 274) = 5.60, p < .001. As can be seen in Table 4, the proportion of know relative to remember ratings was higher for the new items than for the source items, but this did not interact with group, as the three-way interaction was not significant, F(4, 274) = 0.74, p > .5. The main effect of response type achieved significance, F(1, 137) = 62.64, p < .001, indicating that overall know ratings were more often made than remember ratings. The two-way Response Type × Group interaction did not reach significance, F(1, 137) = 1.05, p > .3, suggesting that participants with schizophrenia did not differ from those in the control groups on the proportion of remember and know judgments made. Only the Group × Item Type analysis yielded a significant result, F(4, 274) = 6.81, p < .001. Subsequent post hoc tests showed that this was due to higher percentages of remember and know ratings for the source items in the healthy participants relative to participants with schizophrenia or those with OCD or PTSD (at least p < .01).

As expected, confidence associated with remember responses (M = 3.75) was greater than confidence associated with know responses (M = 2.63) as evidenced by a significant effect of response type (remember vs. know), F(1, 137) = 1,001.68, p < .001, which was not moderated by group, F(2, 135) = 0.46, p > .6.

Relationship to Psychopathology

All major neurocognitive indexes (most notably knowledge corruption, source errors, false-negative errors, false-positive errors, bias to commit errors for the different valence types, confi-
dence in errors and correct responses, number of remember and know ratings, confidence in remember and know ratings) were correlated with core psychopathological variables within the schizophrenia group (PANSS syndrome scores, symptom scores for delusions and hallucinations, chlorpromazine equivalent dosage) and across groups (gender, number of hospitalizations). To limit the occurrence of Type I errors for the core correlational analyses (limit the occurrence of Type I errors for the core correlational analyses \( n = 104 \)), a Bonferroni correction was applied. None of the correlations were significant with this conservative threshold or with a more lenient criterion \( (p < .01) \).

To directly address the hypothesis that hallucinations are associated with an externalization bias for self-generated negatively valenced words, the schizophrenia sample was split into two groups: those who hallucinated \( (n = 14) \) and those who did not hallucinate \( (n = 15) \) for 2 participants, presence of hallucinations could not be reliably assessed. No such bias occurred for any of the three valence types.

**Relationship Between Accuracy and the Confidence Gap**

A reduction in the confidence gap (i.e., increased confidence in errors and decreased confidence in correct responses) may be a natural consequence of poor memory performance. To explore the independence of the confidence gap and accuracy, we carried out a series of correlational and means-comparison analyses.

Across groups, differences occurred with respect to the relationship between memory errors and the confidence gap (i.e., difference in mean confidence for errors and correct responses). While for both healthy participants and those with OCD or PTSD, an increased discriminability between correct and incorrect responses in terms of confidence was moderately correlated with fewer errors \( (r = -29, p = .02) \); OCD or PTSD: \( r = -46, p < .001 \), no such relationship occurred for participants with schizophrenia \( (r = -18, p > .3) \), suggesting that performance and confidence are not redundant measures, particularly for the schizophrenia patients. However, the difference between correlations did not achieve significance \( (p > .6) \). Moreover, when the healthy participants were median-split according to performance, both subsamples still showed an increased confidence gap relative to the participants with schizophrenia \( (p \leq .001) \), despite the absence of a performance difference between the low-performing control group and the schizophrenia group \( (p > .9) \). This finding of aberrant metamemory for schizophrenia despite equating on performance echoes that reported above, whereby the participants with OCD or PTSD did not differ from those with schizophrenia in terms of performance, but did differ in terms of confidence.

**Discussion**

This is the first demonstration that individuals with schizophrenia display an aberration in metamemory performance when compared with both healthy participants and those with OCD or PTSD. Although previous neurocognitive research has consistently shown that participants with schizophrenia perform poorly compared with healthy participants on verbal memory tasks (for meta-analyses, see Aleman, Hijman, de Haan, & Kahn, 1999; Heinrichs & Zakzanis, 1998), the majority of studies that incorporated psychiatric controls were, with few exceptions (Egeland et al., 2003), unable to discriminate among groups (Moritz, Heeren, et al., 2001). The strength of the present approach lies in combining performance accuracy with metamemory parameters. Merging the two aspects of cognition produces a response pattern unique to participants with schizophrenia, which is not predicted by either dimension alone (i.e., participants with schizophrenia did not differ from those with OCD or PTSD when assessing accuracy and overall memory confidence was comparable between groups).

In agreement with prior studies by our group (Moritz & Woodward, 2002; Moritz, Woodward, et al., 2004, 2005; Moritz et al., 2003), participants with schizophrenia were less able than controls to discriminate correct and incorrect responses in terms of confidence: Participants with schizophrenia displayed greater confidence in errors while being somewhat less confident in correct
responses (i.e., decreased confidence gap; Moritz et al., 2003). This was observed both for a parameter that expresses the percentage of high-confidence responses for correct and incorrect responses, as well as for mean confidence ratings for correct and incorrect responses. Thus, the investigation of metamemory provides important nonredundant information to characterize the specific neurocognitive profile of schizophrenia (see also Danion et al., 1999).

That the currently reported metamemory impairment is not redundant with the well-known memory decrement in those with schizophrenia was substantiated by a group comparison of healthy participants and participants with schizophrenia who were equated on accuracy. Healthy control participants who were indistinguishable from those with schizophrenia on recognition performance still displayed an increased confidence gap. Moreover, a modest relationship was detected for the confidence gap with memory accuracy in control groups, but not in the group with schizophrenia. Finally, as noted in the Results section, participants with schizophrenia differed from those with OCD or PTSD on metamemory measures, but not on performance measures.

Greater knowledge corruption in schizophrenia was detected for source memory errors relative to all groups and for false-negative errors relative to healthy controls. A secondary analysis showed that knowledge corruption for source memory was mainly due to knowledge corruption on misattributions of computer-generated words to oneself in schizophrenia. This pattern of results was not seen in a previous study (Moritz, Woodward, et al., 2005) and was unpredicted. Allegedly, procedural differences have contributed to this inconsistency. In prior source memory studies, participants were asked to generate a close associate to a word given by the experimenter, thus probably making such responses more distinctive. In contrast, for the present task, the set-up for participant-generated and computer-generated information was almost identical (verbal clues and solutions were visible for participants in both conditions), with the only difference being that the items were named either by the computer or by the participant. Most important, the solutions to the puzzles were fixed, preventing idiosyncratic responses that would have raised the distinctiveness of the response. The mental episodes of computer- versus self-read material may therefore have been less distinguishable for schizophrenia patients than in the previous studies.

While it may seem counterintuitive that participants with schizophrenia displayed knowledge corruption for misattributions of computer-generated words to oneself, there is some new evidence that links this type of source memory error to delusions (Woodward, Menon, Hu, & Keefe, in press). Moreover, it must be kept in mind that although misattributions of self-generated mental episodes to others (as in hallucinations) are frequently reported in schizophrenia, disturbances of ego boundaries are twofold in schizophrenia: Apart from reality-monitoring errors of the latter type, many individuals with schizophrenia also tend to misattribute external events as self-generated or as partly under their control (e.g., thought broadcasting).

Unlike in prior studies of knowledge corruption, the emotional valence of the material was manipulated. This manipulation was included because it has recently been hypothesized that individuals with schizophrenia show enhanced cognitive biases when confronted with emotionally laden material (Bentall & Kaney, 1989; Danion et al., 2003; Phillips, Senior, & David, 2000; Suslow, Roestel, & Arolt, 2003). In particular, it has been put forward that those with auditory hallucinations may have a tendency to externalize self-generated negative items; that is, contents of consciousness that pose a threat to the ego may be externalized (Bentall et al., 1991; Morrison & Haddock, 1997). This theory, however, was not confirmed. Participants with schizophrenia who experienced hallucinations were indistinguishable from those who did not experience hallucinations on performance for self-generated negative items.

With respect to this negative finding, it can be objected that the negative material in our study was not sufficiently self-referent to trigger such auto-protective behavior. However, another study (Morrison & Haddock, 1997) also found no significant between-groups differences for truly self-generated responses. In view of comparable studies that were able (Baker & Morrison, 1998; Bentall et al., 1991; Brebion, Gorman, Amador, Malaspina, & Sharif, 2002; Laroi et al., 2004; Woodward et al., in press) or unable (Brebion, Smith, Gorman, & Amador, 1997; Keefe, Arnold, Bayen, & Harvey, 1999; Moritz & Woodward, 2002; Moritz et al., 2003; Vinogradov et al., 1997) to obtain group differences, the basic hypothesis should not be prematurely abandoned, but more effort needs to be devoted to elucidate possible moderators of this waxing and waning effect. For example, one suggestion is that guessing strategies can mask pure externalization biases (Keefe et al., 2002; Woodward, Menon, et al., in press).

As is outlined in more detail elsewhere (Moritz, Woodward, et al., 2005), we propose that knowledge corruption in schizophrenia is a consequence of a liberal acceptance bias that is seen in both paranoid and nonparanoid individuals with schizophrenia, whereby fragmented and partial information is taken as sufficient evidence to accept a response option (e.g., a strange sound on the phone may be taken as an indication that the telephone line is being tapped). This mechanism leads, under some conditions, to jumping-to-conclusions behavior, particularly when the set of alternative response options is limited (Moritz & Woodward, 2005), as in probabilistic reasoning tasks (Garety et al., 1991; Haq et al., 1988). It is assumed that the premature completion of data collection, reflecting a liberal acceptance threshold, leads to somewhat decreased confidence in correct judgments in comparison with individuals who collect more supportive evidence for a response and thus raise their confidence. To illustrate, confidence ratings are presumably higher in a participant who is able to retrieve multiple, perceptually rich, conscious memory traces, held in his or her inner ear or inner eye (e.g., that a window was slammed when he or she solved the particular item, that they initially had a different solution in mind) than for a participant who merely knows that an item has occurred without vivid perceptual recollection (Yonelinas, 2001). In line with this argument, it has been found that healthy participants make old–new judgments by means of episodic retrieval of specific episodes (Roediger et al., 2001), whereas those with schizophrenia tend to rely on understanding the gist of the situation (Moritz, Woodward, et al., 2004). Likewise, the absence of such retrieval cues for errors will caution healthy individuals and those with OCD or PTSD against high confidence ratings, whereas participants with a more liberal acceptance threshold may deem such scarce evidence sufficient for a subjective high level of confidence. Participants with schizophrenia reported remember experiences more frequently for incorrect responses, which again
may reflect the tendency to make strong inferences on the basis of little evidence.

At first sight, this chain of arguments may seem at variance with the remember–know data of the present study. It was hypothesized that participants with schizophrenia would recognize items less vividly (as assessed with remember–know ratings) and that they would rely more on familiarity than participants in the control groups (Weiss et al., 2002). As the signature of the latter bias, we expected that participants with schizophrenia would give higher confidence ratings for know ratings, indicating that even incomplete evidence is sufficiently trustworthy. Unlike in prior studies (Danion et al., 1999; Huron & Danion, 2002; Huron, Danion, Rizzo, Killofer, & Damiens, 2003), we did not find that participants with schizophrenia made fewer remember judgments in comparison with those in both control groups, and we found no group differences in confidence for know or remember ratings. Several variables must be taken into account to explain these nonsignificant group differences with regard to the remember–know dichotomy. Most important, prior research in schizophrenia administered only remember–know or confidence ratings. The direct succession of the two metacognitive ratings, as in the present study, may have influenced remember–know judgments, which always followed the confidence ratings. In accordance with this, some participants expressed that they found the two ratings rather redundant. Although it was explained to participants that the two ratings tap different aspects of memory (e.g., someone can be very confident but still have no vivid recollection of a memory), endorsement of high confidence may have automatically primed a remember response to reconfirm a high-confidence rating and a know rating to express uncertainty. Another important procedural difference between the current study and prior work with the remember–know dichotomy is that these studies utilized mostly recognition (Huron et al., 1995, 2003) and not source memory tasks (however, see Danion et al., 1999). In addition, we did not ask participants to provide evidence for their ratings (e.g., Huron et al., 2003). With respect to the nonsignificant differences on remember ratings, some evidence shows that for individuals with schizophrenia, a decrement of conscious recollection is more pronounced for nonverbal than for verbal learning material (Danion et al., 1999; Huron et al., 2003).

It should also be taken into account that the cognitive operations that underlie categorization of responses into remember–know judgments may vary across samples. For example, participants with schizophrenia could have a more lenient criterion for judging a memory as vivid, whereas healthy participants may require more perceptual details to make such a judgment (e.g., recalling that it took a long time to solve a particular puzzle or that table was written as tabelia). However, Huron et al. (2003) recently found that participants with schizophrenia did not differ from healthy controls in the number of visual details that were associated with remember and know ratings; therefore, this explanation is less likely (unfortunately, Huron et al., 2003, did not separate the groups on validity of remember judgments, which might have revealed a different pattern of bottom-up/perceptual support for correct vs. incorrect remember ratings between groups).

Calculation of signal detection parameters on confidence judgments may prove useful to further validate the claim that individuals with schizophrenia have a liberal acceptance bias for high-confidence responses. However, the present study did not allow computation of confidence-based parameters, because this requires separate itemwise confidence ratings for each source (Slotnick, Klein, Dodson, & Shimamura, 2000). Conventionally computed signal detection parameters indicated that participants with schizophrenia were less sensitive to the new–old distinction than were healthy participants (reduced d’), and that participants with schizophrenia may have countered this by avoidance of committing false positives (increased β). However, neither of the two parameters discriminated participants with schizophrenia from those in the control groups.

Another interesting result was that the participants with OCD did not display decreased confidence in their responses. Whereas this finding may seem counterintuitive, it is in agreement with newer findings showing that decreased memory confidence is not ubiquitous in those with OCD (Cabrera, McNally, & Savage, 2001; Radomsky, Rachman, & Hammond, 2001; van den Hout & Kindt, 2003) and may occur only when perceived responsibility is inflated.

Future research may focus on the investigation of neurocognitive rehabilitation programs that involve metacognitive skill training (Koren, Seidman, Poyurovsky, Goldsmith, & Klein, 2004; Moritz, Burlon, & Woodward, 2005). Individuals with schizophrenia may learn to be more aware of common metacognitive errors (e.g., reliance on mere familiarity may facilitate errors in the false memory paradigm, Roediger & McDermott, 1995), may be cautioned about trusting partial information, and may learn ways to confirm hypotheses in a more stringent manner. Similarly, it is important for those with schizophrenia to understand the fallibility of human recollection and to understand that memory information is most reliable when supported by conscious recollection (i.e., remember quality). When such information is not available, judgments should be performed with caution, and additional evidence (e.g., anchor persons such as close relatives) should be consulted for confirmation.

References


Received September 10, 2004
Revision received May 7, 2005
Accepted May 18, 2005