

The Contribution of a Cognitive Bias Against Disconfirmatory Evidence (BADE) to Delusions in Schizophrenia

TODD S. WOODWARD,^{1,2} STEFFEN MORITZ,³
CARRIE CUTTLER,⁴ AND JENNIFER C. WHITMAN²

¹Department of Research, Riverview Hospital, Coquitlam, Canada

²Department of Psychology, Simon Fraser University, Burnaby, Canada

³University Hospital Hamburg-Eppendorf, Hamburg, Germany

⁴Department of Psychology, University of British Columbia, Vancouver, Canada

A neuropsychological paradigm is introduced that provides a measure of a bias against disconfirmatory evidence (BADE), and its correspondence with delusions in people with schizophrenia and schizoaffective disorder was investigated. Fifty-two patients diagnosed with schizophrenia or schizoaffective disorder (36 were acutely delusional) and 24 healthy control participants were presented with delusion-neutral pictures in each trial, and were asked to rate the plausibility of four written interpretations of the scenario depicted by that picture. Subsequently, new pictures that provided background information about the depicted scenario were successively presented, and participants were requested to adjust their ratings, taking into account this new information. Two of the interpretations appeared tenable initially but ultimately proved to be implausible, one appeared untenable initially but eventually proved to be plausible, and one appeared untenable at all stages. A BADE was observed for delusional compared to non-delusional patients, as well as for all patients compared to controls. In addition, regardless of symptom profile, patients were more accepting of implausible interpretations than controls. The present work suggests that deficits in reasoning may contribute to the maintenance of delusions via an impairment in the processing of disconfirmatory evidence.

Delusions are defined as fixed false beliefs not amenable to contrary evidence, and are hallmark symptoms of schizophrenia spectrum disorders. Although it was initially proposed that delusional thought cannot be explained by a pathology of reasoning (Maher, 1988, p. 22), recent work has revealed a number of aberrations in logical thinking that are presumably relevant to our understanding of delusions (for a review see Garety & Freeman, 1999).

Received 10 June 2004; accepted 13 January 2005.

The authors thank Jessica Bristowe, Jannine Laseleta, and Tonya Kragelj for assistance with data collection and data management, and Drs. Karin Christensen, Mahesh Menon, Elton Ngan, Eric Strachan, and Ivan Torres for comments and discussions that shaped this work. This research was supported by post-doctoral fellowships from the Mind Foundation of BC, the Canadian Psychiatric Research Foundation, and the Canadian Institutes of Health Research to TSW, scholarships from the German Academic Exchange Program (DAAD) to SM, and a grant from the Riverview Hospital Academic Steering Committee to TSW.

Address correspondence to Todd S. Woodward, Room 105, Department of Research, Administration Building, Riverview Hospital, 2601 Lougheed Highway, Coquitlam, B.C., V3C 4J2. E-mail: toddswoodward@gmail.com

In the present work we consider deficits in reasoning that may contribute to the maintenance of delusions by focusing on processing of disconfirmatory evidence.

According to Freeman, Garety and colleagues (Freeman, Garety, Kuipers, Fowler, & Bebbington, 2002; Garety, Kuipers, Fowler, Freeman, & Bebbington, 2001), the cognitive operations underlying delusion maintenance can be divided into two mechanisms: obtaining confirmatory evidence (e.g., delusion congruent attentional and memory biases) and discarding disconfirmatory evidence. These authors list a number of cognitive biases that have been put forward that are particularly informative with respect to obtaining confirmatory evidence. For example, a *jumping to conclusions* bias may limit the amount of data gathered (e.g., Garety, Hemsley, & Wessely, 1991), an *attributional bias* may lead to externalization of negative events (e.g., Bentall, Kinderman, & Kaney, 1994), a *theory of mind* dysfunction may lead to errors in reading the intentions of other people (e.g., Frith, 1994); in addition, evidence has been presented that continual expectance of threat could lead to the *misperception of social threat*, the confirmation of delusional ideas, and clinical states of paranoia (Green & Phillips, 2004).

All of the cognitive biases suggested for obtaining confirmatory evidence have been demonstrated on delusion-neutral material, and are therefore clearly separable from the symptoms themselves. In contrast, the mechanisms that have been put forward for disregarding disconfirmatory evidence involve explaining away disconfirming evidence by integrating it into the delusional framework. As such, they are not clearly distinct from the symptom. Freeman and colleagues (Freeman et al., 2002; Garety et al., 2001) suggested two such strategies: *safety behaviors* (e.g., the only reason they did not shoot me yesterday was because I avoided watching television) and *incorporation* (e.g., the only reason they did not shoot me yesterday was because they want me to suffer another day). These authors suggest that all classes of delusions (e.g., paranoid, reference, guilt, grandiose), by definition, share the quality of being maintained in the face of disconfirming evidence. The motivation for the present study was to explore whether a cognitive bias exists against disconfirmatory evidence (BADE) that extends beyond the delusion itself, and is also detectable on delusion-neutral material.

To ensure that a BADE effect (a failure integrating disconfirmatory evidence) could not be attributed to a more generalized “integration” deficit, or to mere perseveration of responses, we designed a control condition requiring the integration of confirmatory evidence. Thus, we compared delusional and non-delusional schizophrenic patients on their ability to integrate confirmatory and disconfirmatory evidence when interpreting neutral scenes. We expected a difference between the two groups when processing disconfirmatory evidence, but not for the control condition involving confirmatory evidence. In addition, we expected to replicate our previous observation (Moritz & Woodward, 2004) of liberal acceptance of implausible interpretations, regardless of the presence of delusions. The signature of a liberal acceptance bias is that all patients with schizophrenia, regardless of symptom profile, overrate the plausibility of absurd hypotheses, because less evidence is deemed necessary to accept a response option.

Methods

Participants

Fifty-two patients diagnosed with schizophrenia or schizoaffective disorder, according to DSM-IV criteria, were recruited from Riverview Hospital and the Forensic Psychiatric Services Commission, Port Coquitlam, BC, Canada. Thirty-six of these patients were

considered actively delusional, and 16 were not currently experiencing delusions (determined by the symptom ratings methodology described below). The actively delusional group contained the following diagnoses: paranoid ($n = 19$), undifferentiated ($n = 11$), and schizoaffective ($n = 6$), and were actively experiencing the following delusions: paranoid ($n = 26$), grandiose ($n = 19$), Schneiderian ($n = 18$), and guilt ($n = 3$). The currently non-delusional group contained the following diagnoses: paranoid ($n = 10$), undifferentiated ($n = 4$), and schizoaffective ($n = 2$). Of the 16 non-delusional patients, 14 displayed delusions in the past, as determined by a comprehensive chart review. Twenty-four healthy control participants, consisting of Riverview Hospital staff were also recruited for the study, and were comparable to the patient groups on age and social status (determined using a modified version of the Hollingshead Index of Social Position; Hollingshead & Redlich, 1958).

The Hollingshead measure of social status is based on the highest parental occupation and education level. We elected to match patients to controls on this variable, as opposed to current IQ, because matching on current IQ is likely to lead to a failure in matching on a host of other variables related to social experiences during development. In addition, compromised intellectual abilities are an important aspect of schizophrenia, and if this illness is to be fully understood, matching for IQ may be counterproductive.

A comparison of demographic information is presented in Table 1. Participants were excluded if their IQ was less than 70, if they had a history of primary or acquired brain damage (e.g., stroke, encephalitis) or traumatic head injury (e.g., with a loss of consciousness for more than 10 minutes), or if they were HIV positive. For all but four of the participants their primary language was English. Eyesight for all participants (assessed corrected and bilaterally) was 20/40 or better. One-way analyses of variance (ANOVAs) and follow-up t-tests demonstrated no difference between the delusional and non-delusional groups on age, education, gender, IQ (as assessed by the QUICK test; Ammons & Ammons, 1962), or social status. In addition, no differences between the patient groups and the healthy control group were present for age or social status. However, as expected, matching by social status led to significant differences between the patient and healthy control groups in IQ, $F(2,70) = 11.71$, $p < 0.001$, and education $F(2,70) = 10.80$, $p < 0.001$. The patient and healthy control groups also differed in gender makeup, $\chi^2(2) = 14.82$, $p < .001$ (see Table 1). In the results we include a section devoted to studying the impact of these potential confounds. At the time of testing, all patients but one were receiving atypical neuroleptic medication (chlorpromazine equivalent dosage in mg: $M = 658.43$, $SD = 488.94$) (Bezchlibnyk-Butler & Jeffries, 2000).

Materials

BADE Task. Initially ten scenarios were used in this study, based on the following WAIS-III (Wechsler, 1997) picture arrangement stories: CAP, BAKE, OPENS, CHASE, CLEAN, HUNT, SAMUEL, LUNCH, CHOIR and SHARK. From each WAIS-III story, only three pictures were presented (not necessarily in the order specified in the WAIS-III manual), and some pictures were edited in order to optimize the fit of the interpretations to the scenarios (see Table 2 for a full description of modifications and presentation order).

For each scenario, four possible interpretations were created: one *true* interpretation (a tenable interpretation of the overall sequence), two *lure* interpretations, designed to appear reasonable at first, but to become increasingly untenable as more evidence regarding the true content of the scenario was presented, and one *absurd* interpretation that was obviously implausible at any point in the trial sequence. For example, the CHASE

Table 1
Sociodemographic characteristics of the samples

	Delusional Patients (D)	Non-delusional Patients (ND)	Healthy controls (C)	ANOVA; Chi-square	Group comparisons
Age	38.08 (9.89)	38.06 (9.11)	39.46(11.50)	NS	
Sex	8 F, 28 M	5 F, 11 M	17 F, 7 M	p. < .001	D,ND ≠ C
Social status	34.16(16.54)	38.50 (11.89)	37.38(16.17)	NS	
Education	11.86 (2.47)	11.50 (1.21)	15.13 (3.12)	p. < .001	D,ND < C
IQ estimate ¹	98.53 (8.61)	99.44 (9.25)	109.62 (8.28)	p. < .001	D,ND < C
Length of illness	17.75 (9.72)	15.69 (8.40)	n/a	n/a	p = .47
Delusions	3.47 (.50)	0.69 (0.79)	n/a	n/a	p < .001
Hallucinations	1.81 (1.81)	0.25 (0.68)	n/a	n/a	p < .01
Formal thought disorder	1.39 (1.17)	0.38 (0.61)	n/a	n/a	p < .01
Flat affect	0.72 (1.00)	1.25 (1.52)	n/a	n/a	p = .15
Poverty of speech	0.22 (0.72)	1.00 (1.36)	n/a	n/a	p < .01
Underactivity	0.53 (0.94)	0.69 (0.70)	n/a	n/a	p = .55

Note. ¹Quick Test (Ammons & Ammons, 1962).

Table 2
Table of modifications to WAIS-III pictures, intended to increase the ambiguity of the scenarios

WAIS-III Picture set name	Order of presentation	Modifications	The true interpretation revealed on (1 st , 2 nd or 3 rd picture)
CAP	3,2,1	None	2 nd
BAKE	3,4,1	None	1 st
OPENS	5,3,1	None	3 rd
CHASE	4,1,2	None	2 nd
CLEAN	1,2,5	None	1 st
HUNT	1,3,4	Prisoner number erased on pictures 1 and 3	3 rd
SAMUEL	2,1,4	None	3 rd
LUNCH	1,5,4	Apple erased on all pictures	1 st
CHOIR	2,1,3	None	2 nd
SHARK ^a	4,2,5	Real shark erased on picture 4	2 nd

Notes. ^aUsed for practice trial only.

scenario was designed as follows: first, a picture of a man with a small dog at his feet, looking over a fence at a large barking dog was presented along with the following interpretations (presented without the “lure”, “true”, and “absurd” text given here in brackets): (1) “The man has just escaped from the barking dog” (true); (2) “The man is looking at his neighbour’s barking dog” (lure); (3) “The man has just built a fence for his dog” (lure); and (4) “The man is shopping for guard dogs” (absurd). At this point,

participants typically endorse interpretations 2 (looking at neighbor's dog) and/or 3 (built a fence), that is to say, they endorse the lures. Following this, a picture of the same man jumping over the fence, with the dog barking close to the man's heels and pawing the fence was presented, and the participants were told that this is the situation that immediately preceded the scenario depicted in the picture that they just interpreted, and were asked if they would like to revise their original ratings. At this point, the lures are typically down-rated, and the true interpretation is up-rated. Following this, a third picture was presented depicting a barking dog chasing the running man towards the fence. The participants were told that this situation immediately preceded the scenario depicted in the two pictures that they just interpreted, and were asked if they would like to revise their original ratings a final time. At this point, the lures are typically down-rated even further, and the true interpretation up-rated even further. Ratings for the absurd interpretations are not typically affected by the presentation of the second and third pictures. This procedure is described in more detail below.

The SHARK scenario was designated as a practice trial. Of the nine remaining scenarios, three were designed as *revealed-on-first*, three as *revealed-on-second*, and three as *revealed-on-third*. For revealed-on-first scenarios, the true interpretation was obvious when the first picture was presented, and the next two pictures reinforced this interpretation. For revealed-on-second scenarios, the true interpretation was obvious when two pictures were visible. For revealed-on-third trials, the true interpretation was obvious only when all three pictures were visible. Revealed-on-second and revealed-on-third trials were designed to elicit a measurement of integration of confirmatory and disconfirmatory evidence. These latter two trial types differed qualitatively in the sense that the to-be-changed ratings were maintained for two pictures on revealed-on-third trials, but for only one picture on revealed-on-second trials. Revealed-on-first trials did not include lure interpretations, so did not require further information to reveal the correct interpretation; therefore they were considered filler items, and were included to avoid predictability of when the true interpretation became obvious.

For revealed-on-second and revealed-on-third scenarios, true interpretations were designed such that they would appear less plausible than lure interpretations prior to presentation of the true-interpretation-revealing picture. This aspect of the task design was central to the task validity; if lure interpretations were considered less plausible than true interpretations on presentation of the first picture, a substantial reduction in power for detecting the integration of disconfirmatory evidence would result, and these lures should be considered invalid.

The presentation order and classification is listed in Table 2. Prior to the main analyses, the distinction between the true, lure and absurd interpretations was confirmed using the observed ratings for all subjects. For all scenarios, once all three scenario cards were displayed, the superiority of the true rating was clearly indicated for all scenarios, and for no trials was a mean rating for a true or lure interpretation lower than an absurd interpretation. As a more thorough validity check, ratings for the revealed-on-second and revealed-on-third trials were inspected to ensure that the lure interpretations were rated higher than the true interpretations initially, but lower once the true interpretation was revealed. This criterion was met for all groups and all interpretations with the exception of the CAP scenario, for which the plausibility of true and lure interpretations were reversed for all three groups on the first scenario card (i.e., the participants were not misled by the lure for this scenario). Therefore, CAP trials were excluded from further data analyses. The full set of final scenarios and interpretations is available by request from the first author.

Each interpretation was printed in 20 point Times New Roman font onto 13 × 8 cm cards, and arrows were situated on the bottom center of each card to facilitate precise placement on the rating strip. Plausibility ratings were made by placing these cards on one of four rating strips which were attached to a 71 cm × 55 cm rating board in evenly spaced horizontal rows. Directly below each rating strip was a measuring tape (71 cm long), with the labels “poor”, “possible”, “good”, and “excellent” at 9 cm, 27 cm, 42 cm, and 60 cm, respectively. In addition, full descriptions of the rating options were placed at the top of the rating board, and were also centered at 9 cm, 27 cm, 42 cm, and 60 cm; these read “does not fit/poor interpretation”, “possible but unlikely”, “good interpretation”, and “excellent”, respectively. The use of four rating strips ensured that the plausibility rating for any given interpretation was not physically constrained by the placement of any other interpretation. In other words, because interpretation cards were placed on separate rating strips, participants had the option of giving any number of interpretations exactly the same plausibility rating.

The purpose of using this rating method, as opposed to the traditional Likert-type scales, was (1) to add a simplifying spatial dimension to the ratings, such that patients could view a left-to-right observable scale for all ratings at all times, thereby facilitating comparison of ratings across interpretations; and (2) to allow minute adjustments to be detected by using a continuous as opposed to a categorical scale. In order to ensure that participants understood that left-to-right corresponded to low-to-high plausibility, the rating scales were clearly labelled at the top and bottom of the rating board, as well as under each rating strip, and this was clearly explained to all participants. Feedback from the participants indicated that this methodology was straightforward and well understood, and our experience suggests that the physical dimension introduced by this method is preferred to the verbal single-number response typically employed by Likert-type scales.

Neuropsychological/Cognitive Measures

Measures indexing integration of confirmatory and disconfirmatory evidence were computed using change scores. More specifically, the disconfirmatory measure was computed as the *decrease* in plausibility ratings from picture #1 to picture #3 for lure interpretations, and the confirmatory measure was computed as the *increase* in plausibility ratings from picture #1 to picture #3 for true interpretations. These confirmatory and disconfirmatory change scores were computed separately for revealed-on-second and revealed-on-third trials. A separate measure, designed to index liberal acceptance, was computed as the mean plausibility rating for absurd interpretations, averaged over all trial types (i.e., averaged over revealed-on-first, revealed-on-second, and revealed-on-third trials).

Symptom Measures

Psychopathology was assessed using the SSPI (Liddle, Ngan, Duffield, Kho, & Warren, 2002). The SSPI is a 20 item scale, which can be completed after a 25 – 30 minute semi-structured interview with 15 direct questions about symptoms. The severity of each item is rated on a scale ranging from 0 to 4. The SSPI is criterion referenced, providing specific examples of behavior which correspond to severity levels for each item. Item 7 from the SSPI was used to quantify the presence/absence of delusions. A rating of 3 (definite delusions, but the delusional beliefs do not have a pervasive influence on thinking or behavior) or 4 (definite delusions which have pervasive influence on thinking and/or influence observable behavior) warranted classification into the delusional group.

Although the SSPI was typically carried out on the day of cognitive testing, the maximum allowable time gap was two weeks.

Procedure

For each trial, patients were presented with a series of three pictures that described a scenario. After the presentation of the first picture (picture #1), participants were read four possible interpretations of the depicted scenario. The interpretations and trials were presented in a random order that differed for each participant. After each interpretation was read, that interpretation card was placed in front of the participant. Once all four cards were placed in front of the participant, they were asked to place them on the rating board, starting with the first interpretation card on the top-most rating strip, and working their way down to the fourth interpretation card on the bottom-most rating strip. An arrow pointer positioned on the bottom of each card determined the plausibility rating, and this position was recorded to the nearest 0.5 cm by the experimenter.

The second picture (picture #2) was then placed to the left of picture #1, and participants were told: "Now to give you more information about what's going on in this picture (experimenter points to picture #1) I am going to show you another picture (experimenter points to picture #2) that shows you what happened just before this (experimenter points to picture #1)." Participants were then required to describe the depicted events, and were corrected as required until it was clear that the sequence of events was fully understood by the participant. The participant was then told: "Now given this new information (experimenter points to picture #2) about what happened just before this (experimenter points to picture #1) adjust your previous ratings of this picture (experimenter points to picture #1) if you wish. Just to be clear: the interpretations you are rating are still of this picture (experimenter points to picture #1). The other picture just gives you more background about what's going on." After recording any adjustments in plausibility ratings, the experimenter placed down the third and final picture (picture #3) to the left of the picture #2. At this point, the experimenter said "So it makes a story. This happened first (experimenter points to picture #3) then this happened (experimenter points to picture #2) then this happened (experimenter points to picture #1). Do you understand? Now briefly tell me the events from beginning to end." The participant was corrected if they misunderstood the chronological order or the story. Once the experimenter was satisfied that the participants understood the chronological order and the story, the participants' summary was recorded, and the participants were asked to make any final adjustments in their plausibility ratings. In total, each of the four interpretations was rated three times, once after each picture was presented. The total duration of the test session was approximately 45 minutes.

Results

The plausibility ratings were analyzed by one-way Analyses of Variance and Analysis of Covariance (ANOVAs and ANCOVAs). Group differences were analyzed with either "reverse Helmert" contrasts, or "repeated" contrasts, with the groups ordered as follows: delusional patients, non-delusional patients, and healthy controls. With this group ordering, reverse Helmert contrasts provide one-degree-of-freedom tests for a difference between the two patient groups (contrast: $-1\ 1\ 0$), and for the difference between all patients (averaging the two patient groups) and healthy controls ($-.5\ -.5\ 1$). If no difference between the patient groups was observed, reverse Helmert contrasts were reported.

Table 3

Covariate-adjusted mean *decreases* (for Lure Interpretations; BADE) and *increases* (for True Interpretations), in cm, presented as a function of group (standard errors in brackets). Superscripts indicate significant mean differences assessed using post-hoc t-tests (least significant difference, two tailed)

Effect	Group		
	Delusional	Non-delusional	Healthy controls
Disconfirmatory evidence (Lure)			
Revealed-on-second	9.96 (2.15) ^{§§}	14.49 (3.19) [†]	23.68 (2.64) ^{†§§}
Revealed-on-third	15.37 (1.73) ^{*§§§}	21.78 (2.59) [*]	27.26 (2.13) ^{§§§}
Confirmatory evidence (True)			
Revealed-on-second	26.53 (1.86)	23.79 (2.78)	29.58 (2.26)
Revealed-on-third	26.63 (1.91) ^{§§}	28.56 (2.84)	34.15 (2.33) ^{§§}

Note. Means are adjusted for the plausibility ratings prior to exposure to the true interpretation. * $p < .05$, delusional patients vs. non-delusional patients; [†] $p < .05$, non-delusional patients vs. controls; ^{§§} $p < .01$, delusional patients vs. controls; ^{§§§} $p < .001$, delusional patients vs. controls.

However, if a difference between the two patient groups was observed, repeated contrasts were used. Repeated contrasts allowed one-degree-of-freedom tests of the difference between delusional and non-delusional patient groups (1 -1 0), and the difference between the non-delusional patient group and the healthy control group (0 1 -1). These a priori contrasts were used in an attempt to constrain reports of statistical significance to the theoretically most important effects. All group differences were assessed using two-tailed significance tests.

A reluctance to integrate disconfirmatory evidence, and an enhanced willingness to integrate confirmatory evidence, may be natural consequences of holding a strong belief (Hemsley, 1988, p. 118; Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Nickerson, 1998). In order to control for the potential impact of the initial belief on changes in that belief, plausibility ratings prior to exposure of the true interpretation (i.e., prior to exposure to the picture that made the correct interpretation of the depicted scenario obvious) were used as covariates for all analyses of rating changes in lure and true interpretations. Thus, true and lure interpretation rating changes were analyzed by way of one-way ANCOVAs with Group as a between subjects factor, and initial plausibility ratings as a covariate. However, it should be noted that no significant group differences were observed on the covariate (viz., initial plausibility ratings for true or lure interpretations; $p > .1$ for all). Covariate-adjusted mean decreases and increases for lure and true interpretations, respectively, are presented in Table 3. Absurd interpretation ratings were analyzed by way of a one-way ANOVA with Group as a between subjects factor, collapsed over all trial types.

Due to qualitative differences between revealed-on-second and revealed-on-third trials (viz., to-be-changed ratings were maintained for two pictures on revealed-on-third trials, but for only one picture on revealed-on-second trials), revealed-on-second and revealed-on-third trials were analyzed separately for lure and true interpretations.

Integration of Disconfirmatory Evidence (Lure Interpretations)

For revealed-on-second trials, a significant effect of Group was observed, $F(2, 72) = 7.93$, $p < .001$. Although the direction of the means corresponded to our predictions, reverse

Helmert contrasts on the change score for lure interpretations from picture #1 to picture #3 failed to reach significance for the comparison of patient groups, $t(72) = 1.17, p = .25$. This analysis was significant when comparing all patients to the healthy controls, $t(72) = 3.49, p < .001$, suggesting that healthy controls decreased their ratings for lures more than patients for revealed-on-second trials.

As for revealed-on-second trials, for revealed-on-third trials, a significant effect of Group was observed, $F(2, 72) = 7.67, p < .001$. In contrast to revealed-on-second trials, and in correspondence with our predictions, repeated contrasts on the change score for lure interpretations from picture #1 to picture #3 reached significance for the comparison of patient groups, $t(72) = -2.05, p < .05$, but not for the comparison of the nondelusional patients with the healthy controls, $t(72) = -1.63, p = .11$. This suggests that, compared to nondelusional patients, delusional patients adhere to their initial judgments in face of disconfirmatory evidence for revealed-on-third trials.

Integration of Confirmatory Evidence (True Interpretations)

For revealed-on-second trials, the effect of Group was not significant, $F(2, 72) = 1.37, p = .26$. For revealed-on-third trials, a significant effect of Group was observed, $F(2, 72) = 3.11, p = .05$. Repeated contrasts on the change score did not reach significance for the comparison of patient groups, $t(72) = 0.56, p = .58$, but did reach significance when comparing all patients with the healthy controls, $t(72) = 2.26, p < .05$.

Liberal Acceptance of Implausible Information (Absurd Interpretations)

Plausibility ratings for absurd interpretations, averaged over all trials (i.e., revealed-on-first, revealed-on-second, and revealed-on-third), revealed a significant effect of Group, $F(2, 73) = 9.28, p < .001$. Healthy controls, at each stage of the experiment, rated absurd interpretations lower ($M = 10.03$ cm) than either delusional ($M = 19.19$ cm) or nondelusional ($M = 16.09$ cm) participants. In agreement with our previous study (Moritz & Woodward, 2004), reverse Helmert contrasts demonstrated that plausibility ratings averaged over trials did not differ significantly between delusional and nondelusional patients, $t(73) = -1.23, p = .21$, but differed significantly between all patients and healthy controls, $t(73) = -3.71, p < .001$.

Specificity of BADE

To determine that the cognitive variables of interest are specifically related to delusions and not the other positive symptoms that differed across groups (see Table 1), three ordinary least squares (OLS) regression analyses were performed. The criterion variables were the average change score for integration of disconfirmatory evidence (lure interpretations), the average change score for integration of confirmatory evidence (true interpretations), and liberal acceptance of implausible information rating score (absurd interpretations). For all three analyses, the predictor variables were delusions (SSPI item 7), hallucinations (SSPI item 8) and formal thought disorder (SSPI item 17). Due to the expected inverse relationship between symptoms and normal cognitive responding, one-tailed significance tests were used. The betas that results from OLS regression reflect only variance that is unique to each predictor (Pedhazur, 1982), simplifying interpretational difficulties that arise due to overlapping positive symptoms. The analysis of lure interpretations resulted in significance for delusions, $t(47) = -1.70, p < .05$, but not for hallucinations, $t(47) = -0.52, p = .30$, or thought disorder, $t(47) = 1.24, p = .89$, supporting the specificity of the BADE

effect. The analysis of true interpretations resulted in marginal significance for thought disorder, $t(47) = -1.68, p = .05$, but nonsignificance for hallucinations, $t(47) = -0.71, p = .24$, and delusions, $t(47) = 0.08, p = .53$. Finally, the analysis of absurd interpretations resulted in nonsignificance for delusions, $t(47) = 0.00, p = .50$, hallucinations, $t(47) = -0.02, p = .49$, and thought disorder, $t(47) = -1.03, p = .16$.

To ensure that the BADE effect was not redundant with IQ, we computed correlations between IQ and all effects of interest (i.e., Integration of Disconfirmatory Evidence, Integration of Confirmatory Evidence, and Liberal Acceptance of Implausible Information). None of these correlations reached significance with one-tailed tests (all $p > .5$) for the patient or control groups.

Discussion

Although disregarding disconfirmatory evidence is an important aspect of delusion maintenance in schizophrenia, it has never been firmly established whether patients display a generalized bias against disconfirmatory evidence (BADE); that is whether a BADE is restricted to specific delusions, or whether such a bias extends to other interactions with the environment. In the present study, participants were presented with delusion-neutral pictured scenes, and were asked to rate the plausibility of each of four written interpretations. For each critical scene, two of these interpretations appeared tenable initially, but eventually proved to be inaccurate once clarifying background information was presented. Our hypothesis that relative to nondelusional patients, delusional patients would show a decreased change in plausibility ratings was supported for revealed-on-third scenarios. Interpretation of the BADE finding was strengthened by the observation that on these same scenarios, relative to non-delusional patients, delusional patients were *not* impaired in their ability to integrate *confirmatory* evidence into their plausibility ratings. This combination of results suggests that a BADE may contribute to delusion maintenance in schizophrenia, and that differences between delusional and nondelusional patients on BADE are unlikely to reflect a more generalized deficit in integration of information.

The failure of the BADE effect to differentiate significantly between delusional and nondelusional patients on the revealed-on-second trials may be due to qualitative differences between revealed-on-second and revealed-on-third trials. For example, on revealed-on-third trials, the to-be-changed ratings were maintained for two pictures, but were maintained for only one trial on revealed-on-second trials. This suggests that association between BADE and delusions may be enhanced for beliefs that have been strengthened by experience, or reinforced by additional evidence. The absence of a patient group difference for the BADE effect on revealed-on-second trials may also be attributable to the fact that one of the revealed-on-second trials (CAP) was excluded due to validity concerns, possibly rendering this condition less powerful in terms of effect detection. A related point that is reinforced by the trend towards the predicted pattern of results (see Table 3, revealed-on-second trials for lure interpretations) is that because the nondelusional group was relatively small ($n = 16$), power for finding medium sized effects was limited.

It is important to consider the nature of the cognitive deficit that may lead to a BADE. Candidate levels for the impairment are attention, perception, recognition/attachment of meaning to a percept (i.e., determination of disconfirmatory vs. confirmatory information), or reasoning (i.e., integration of disconfirmatory and confirmatory information). The first two stages (attention and perception) can be considered early information processing, and the latter two stages (recognition of meaning and integration) can be considered late information processing.

A deficit at the perceptual level, with perception defined as construction of a conscious percept from sense data *prior to accessing meaning* (Shallice, 1988, p. 188), is unlikely, as the distinction between confirmatory and disconfirmatory evidence is not “pre-meaning”, as are lines and curves to the perception of objects. A deficit at the level of attention to disconfirmatory evidence does not fit the data well either, because, under these circumstances, a reciprocal relationship between attention to confirmatory and disconfirmatory evidence would be expected, such that reduced attention to disconfirmatory evidence would lead to increased attention to confirmatory evidence; this reciprocity was not observed (see Table 3, Lure vs. True interpretations). In addition, all patients were required to demonstrate a full understanding of the presented story and its chronological order (their spoken summary was recorded after the completion of each trial sequence) prior to making their final ratings, thus minimizing the impact of attentional biases on ratings.

We conceptualize BADE as occurring at the later stages of information processing, either at the level of recognizing the disconfirmatory meaning of information, or at the level of using this meaning to reason. Referring to the early Bayesian reasoning literature (Fischoff & Beyth-Marom, 1983; Hemsley & Garety, 1986), the stages of information processing that we are proposing as candidates for the BADE effect are (1) assessing the likelihood ratio (recognition of the disconfirmatory nature of the information), or (2) aggregation (integration of the disconfirmatory evidence for reasoning). For example, in the CHASE scenario described previously, deluded patients may not recognize that a man running away from a dog is disconfirmatory in relation to the “just built a fence for his dog” interpretation (impaired recognition of meaning), or they may recognize this, but not integrate this information in their reasoning and ratings. Our current data do not allow us to distinguish between these two possibilities, although we can conclude that integration of *any type* of information is not the impairment, due to the successful integration of confirmatory evidence.

The higher overall plausibility ratings of absurd interpretations for all patients (regardless of symptom profile) replicates the previously reported liberal acceptance effect (Moritz & Woodward, 2004). The liberal acceptance account holds that a core disturbance associated with schizophrenia is that initially more explanations are taken into consideration when interpreting complex events, whereas healthy participants are more selective, and rule out certain hypotheses more quickly. With respect to this finding, it is important to clarify that patients’ ratings for True interpretations did not differ from those of controls, but that this effect was restricted to scenarios that should be ruled out by “common sense”, or prior knowledge. As has been presented in more detail elsewhere (Moritz & Woodward, *in press*), the liberal acceptance account is a variant of the jumping to conclusions account put forward by Garety and coworkers (e.g., Garety, Hemsley, & Wessely, 1991).

A limitation of this study was that the patient groups differed on symptomatology other than delusions, such that the delusional group displayed more positive symptoms, and the nondelusional group more negative symptoms (see Table 1). This raises concerns that the group differences attributable to delusions may actually be attributable to other symptoms (such as hallucinations), or to severity of illness. Given that performance in crucial control conditions, general cognitive ability, and length of illness did not differ between the delusional and nondelusional groups, it seems unlikely that the delusion-related group difference on BADE would be attributable to differences in severity of illness. Moreover, the regression analyses presented as evidence of specificity suggests that BADE is associated with variance in delusions that does not overlap with variance in hallucinations and thought disorder.

An unexpected finding that resulted from the OLS regression suggested that increases in thought disorder may be associated with decreased integration of confirmatory evidence. A failure to up-rate true interpretations may suggest that the exact correct interpretation of the scenario was not realized. Fast and far spreading semantic activation is thought to underlie formal thought disorder (Moritz, Woodward, Küppers, Lausen & Schickel, 2002), and this may lead to a failure in realizing and associating the semantic concepts necessary for accurately interpreting potentially ambiguous scenarios.

The suggestion that a generalized cognitive bias may be associated with the presence of delusions contrasts with the classical perspective as represented, for example, by Maher (1988, p. 22): "Hence the folk-clinical observation that delusional patients do not readily abandon their theory in the face of critical contradictory evidence does not indicate pathology of reasoning. It merely tells us that deluded patients are like normal people — including scientists — who seem extremely resistant to giving up their preferred theories even in the face of damningly negative evidence." More recently it was suggested out that although cognitive biases are detectable in delusional patients, they do not extend beyond delusional thinking (e.g., Chadwick & Lowe, 1994, p. 363). The preliminary evidence presented here contributes to the literature suggesting that deluded patients may demonstrate generalized cognitive biases that extend beyond delusion-relevant material.

References

- Ammons, R.B., & Ammons, C.H. (1962). *The Quick Test (QT)*. Missoula (MT): Psychological Test Specialists.
- Bentall, R.P., Kinderman, P., & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: Towards a model of persecutory delusions. *Behaviour Research and Therapy*, *32*, 331–341.
- Bezchlibnyk-Butler, K.Z., & Jeffries, J.J. (Eds.) (2000). *Clinical handbook of psychotropic drugs* (10 ed.). Seattle: Hogrefre & Huber.
- Chadwick, P.D.J., & Lowe, C.F. (1994). A cognitive approach to measuring and modifying delusions. *Behaviour Research and Therapy*, *32*, 355–367.
- Fischoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, *90*, 239–260.
- Freeman, D., Garety, P.A., Kuipers, E., Fowler, D., & Bebbington, P.E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology*, *41*, 331–347.
- Frith, C. (1994). Theory of mind in schizophrenia. In A.S. David, & J.C. Cutting (Eds.), *The neuropsychology of schizophrenia* (pp. 147–161). Hillside, NJ: Erlbaum.
- Garety, P.A., & Freeman, D. (1999). Cognitive approaches to delusions: A critical review of the evidence. *British Journal of Clinical Psychology*, *38*, 113–154.
- Garety, P.A., Hemsley, D.R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients: Biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease*, *179*, 194–201.
- Garety, P.A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P.E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine*, *31*, 189–195.
- Green, M., & Phillips, M.L. (2004). Social threat perception and the evolution of paranoia. *Neuroscience and Biobehavioral Reviews*, *28*, 333–342.
- Hemsley, D.R. (1988). Psychological models of schizophrenia. In E. Millner & P. Cooper (Eds.), *Adult abnormal psychology* (pp. 101–127). London: Churchill Livingstone.
- Hemsley, D.R., & Garety, P.A. (1986). The formation of maintenance of delusions: A Bayesian analysis. *British Journal of Psychiatry*, *149*, 51–56.
- Hollingshead, A.B., & Redlich, F.C. (1958). *Social class and mental illness*. New York: John Wiley and Sons.

- Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology, 80*, 557–571.
- Liddle, P.F., Ngan, E.T.N., Duffield, G., Kho, K., & Warren, A.J. (2002). The signs and symptoms of psychotic illness: A rating scale. *British Journal of Psychiatry, 180*, 45–50.
- Maher, B.A. (1988). Anomalous experience and delusional thinking: The logic of explanations. In T.F. Oltmanns & B.A. Maher (Eds.), *Delusional Beliefs* (pp. 15–33). New York: John Wiley.
- Moritz, S., & Woodward, T.S. (2004). Plausibility judgement in schizophrenic patients: Evidence for a liberal acceptance bias. *German Journal of Psychiatry, 7*, 66–74.
- Moritz, S., & Woodward, T.S. (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *British Journal of Clinical Psychology, 44*(2), 193–207.
- Moritz, S., Woodward, T.S., Küppers, D., Lausen, A., & Schickel, M. (2002). Increased automatic spreading of activation in thought disordered schizophrenic patients. *Schizophrenia Research, 59*(2–3), 181–186.
- Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175–220.
- Pedhazur, E.J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (Second ed.). Toronto: Holt, Rinehart and Winston.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Wechsler, D. (1997). *WAIS – III. Wechsler Adult Intelligence Scale* (Third ed.). San Antonio: The Psychological Corporation.

Copyright of Journal of Clinical & Experimental Neuropsychology is the property of Psychology Press (UK) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.