

**Complete title:** Impaired Evidence Integration and Delusions in Schizophrenia

**Running title:** Evidence Integration

William J. Speechley<sup>1</sup>

Steffen Moritz<sup>2</sup>

Elton T.-C. Ngan<sup>1</sup>

Todd S. Woodward<sup>1,3</sup>

(1) Department of Psychiatry, University of British Columbia, 2C1- 2255 Wesbrook Mall, Vancouver B.C., V6T 2A1.(e-mail: [willspeechley@gmail.com](mailto:willspeechley@gmail.com), [eltonngan@gmail.com](mailto:eltonngan@gmail.com))

(2) Universitätsklinikum Hamburg-Eppendorf, Klinik für Psychiatrie und Psychotherapie, Arbeitsgruppe Klinische Neuropsychologie, Martinstraße 52, D-20246 Hamburg. (e-mail: [moritz@uke.uni-hamburg.de](mailto:moritz@uke.uni-hamburg.de))

(3) BC Mental Health and Addiction Research Institute, Translational Research Building, 3rd Floor, 938 W. 28th Avenue, Vancouver, British Columbia, Canada, V5Z 4H4. (e-mail: [todd.woodward@ubc.ca](mailto:todd.woodward@ubc.ca))

**Corresponding author:** Todd S. Woodward, Ph. D., Assistant Professor, Room A3-A116, BC Mental Health & Addictions Research Institute, Translational Research Building, 3rd Floor, 938 W. 28th Avenue, Vancouver, British Columbia, Canada, V5Z 4H4, Tel: 604-875-2000 x4724, Fax: 604-875-3871, e-mail: [todd.woodward@ubc.ca](mailto:todd.woodward@ubc.ca); website:

[http://www3.telus.net/Todd\\_S\\_Woodward](http://www3.telus.net/Todd_S_Woodward)

## **Abstract**

A bias against disconfirmatory evidence (BADE) appears to be related to delusions in schizophrenia. However, preliminary studies have either not used the most comprehensive version of the BADE task, not included a psychiatric control group, and/or have used difference score methodology instead of analyzing all available measures. In the current study a comprehensive version of the BADE task was administered to people with schizophrenia, bipolar disorder and a healthy control group. The BADE task required rating four interpretations of delusion-neutral scenarios three times (in sequence) as increasingly disambiguating information was presented. A principal component analysis (PCA) carried out on all measures determined that two independent cognitive processes appear to combine to determine all responses on the BADE task: Integration of Evidence and Conservatism, with only the former discriminating between the severely delusional schizophrenia group and all other groups. Thus, Integration of Evidence appears to be functioning sub-optimally in severely delusional schizophrenia patients, resulting in a bias against disconfirmatory evidence (BADE). The cognitive process theorized to be underlying this effect is the hypersalience of evidence-hypothesis matches.

**Key Words:** DECISION MAKING, PRINCIPAL COMPONENT ANALYSIS, PSYCHOSIS, REASONING, PROBABILISTIC REASONING, JUMPING TO CONCLUSIONS

## **Impaired Evidence Integration and Delusions in Schizophrenia**

A defining feature of delusional beliefs in schizophrenia is that they are maintained with conviction despite their apparent falsehood. This feature of delusions is captured by the bias against disconfirmatory evidence (BADE; Woodward, Moritz and Chen, 2006; Woodward, Moritz, Cuttler & Whitman, 2006), found that a BADE, which describes a maladaptive insensitivity towards evidence that should cause an individual to re-evaluate their beliefs, exists even for delusion-neutral scenarios and presumably reflects a risk factor for delusion formation. We have developed a family of tasks in which participants are sequentially presented with three statements that increasingly disambiguate a delusion neutral scenario.

In the most comprehensive version of the BADE tasks, participants are required to rate the plausibility of four interpretations of a scenario three times. Three sequential ratings of each interpretation are required because participants are exposed to new evidence about the scenario interpretation after the presentation of each scenario statements. Two of the interpretations are initially plausible explanations that later require revision, and are referred to as *lures*. One is an option that is highly implausible from the outset and remains so, and is referred to as an *absurd* interpretation. Finally, one is initially moderately plausible, but gradually become the most plausible, and is referred to as a *true* interpretation. Participants' willingness to adjust their plausibility ratings for the scenario interpretations as disambiguating evidence accumulates over the three statement presentations is measured. In multiple studies we have found that BADE appears to be a robust effect in schizophrenia (Moritz & Woodward, 2006; Woodward, Moritz, Menon & Klinge, 2008) and that the effect is stronger for delusional patients (Woodward, Moritz and Chen, 2006; Woodward, Moritz, Cuttler & Whitman, 2006).

The BADE task has cycled through a number of experimental iterations, involving picture scenarios (Woodward, Moritz, Cuttler & Whitman, 2006; Moritz & Woodward, 2006), sentence scenarios (Woodward, Moritz and Chen, 2006; Woodward, Moritz, Menon & Klinge, 2008), variations in the number of scenarios (ranging from 8-30), variations in the number of interpretations for each scenario (ranging from 2-4), and has been administered in English (Woodward, Moritz, Cuttler & Whitman, 2006), Cantonese (Woodward, Moritz and Chen, 2006) and German (Woodward, Moritz, Menon & Klinge, 2008; Moritz & Woodward, 2006). The current, most comprehensive English version of the BADE task consists of 30 scenarios and 4 interpretations per scenario.

To date, our published work utilising the most comprehensive version of the BADE task has been limited to a group of university students screened for schizotypal traits (Buchy, Woodward, & Liotti; 2007; Woodward, Buchy, Moritz, & Liotti, 2007). In those studies we found an association between schizotypy and a BADE. In the current study, our goal was to determine whether or not the BADE effect replicates with clinically significant delusions using the comprehensive version of the BADE task, and to ensure that this effect does not extend to a psychiatric control sample. For a psychiatric control group we selected bipolar disorder due to the similarities between people with a diagnosis of bipolar disorder and people with a diagnosis of schizophrenia with regards to cognitive function, and genetic and environmental susceptibility factors (Barrett, Mulholland, Cooper & Rushe, 2009). Therefore, any aspect of task performance attributable to these factors (or other overlapping characteristics, such as the stigmatization associated with mental illness) should be present for both people with schizophrenia and people with bipolar disorder.

In earlier research the BADE effect was most often operationalized as the degree of reduction in plausibility ratings for “lure” interpretations, which appear plausible initially, but are rendered implausible as incoming evidence clarifies the scenario (Woodward, Moritz, Cuttler & Whitman, 2006; Moritz & Woodward, 2006; Woodward, Moritz and Chen, 2006; Woodward, Moritz, Menon & Klinge, 2008; Buchy, Woodward & Liotti 2007). These conclusions were based on univariate analyses, and therefore did not consider information available from overlap between the BADE items. The use of all available measures can simplify interpretations of test performance by considering overlap and redundancy between test items to determine the nature of the cognitive operations underlying all responses. Previously, using data from a group of university students screened for schizotypal traits, we carried out a multivariate analysis on BADE lure items and a number of traditional neuropsychological measures, determining that only two cognitive operations contributed to the lure rating scores (Woodward, Buchy, Moritz & Liotti, 2007). We labeled these “Initial Belief” (composed of lure ratings after statement one) and “Integration of Disconfirmatory Evidence” (composed of lure ratings following statement three). Thus, all previous BADE theorizing focused on the lure items alone, either using univariate analysis on change scores or multivariate analysis on lure ratings. One of the goals of the current study was to use all items (true, absurd and lure) in a multivariate analysis to allow all ratings to contribute to interpretation of the cognitive underpinnings of the evidence integration necessary in the BADE task.

In the current study our goals were (1) to administer this comprehensive version of the BADE test to a sample of people diagnosed with schizophrenia in order to replicate the association between a BADE and delusions in schizophrenia, (2) to compare the performance of people diagnosed with schizophrenia to a psychiatric control group, and (3) to use all items (true,

absurd and lure) in a multivariate analysis to identify the cognitive underpinnings of the BADE task, and how they differ across patient groups. Based on our analysis of the schizotypy data (Woodward, Buchy, Moritz & Liotti, 2007) we anticipated that two cognitive factors would emerge from all BADE items, and that a diagnosis of schizophrenia with severe delusions would result in the most impairment in evidence integration, with healthy controls showing the least evidence of a BADE, and schizophrenia with mild/no delusions and bipolar disorder falling in between.

## **Methods**

### *Participants*

Fifty people with a diagnosis of schizophrenia and 60 people with a diagnosis of bipolar disorder were recruited from psychiatric hospitals and community health agencies in and around Greater Vancouver, British Columbia, Canada. All diagnoses were made by ward psychiatrists and, if in doubt, confirmed using a neuropsychiatric interview (MINI; Sheehan, Lecrubier and Sheehan, 1998) conducted by trained research staff. Psychopathology was assessed using the Signs and Symptoms of Psychotic Illness scale (SSPI; Liddle, Ngan, Duffield, Kho & Warren, 2002), a schedule gauging symptom severity using 20 symptom items scored 0–4. As with our previous work (Speechley, Whitman & Woodward, 2010) the SSPI Item 7 (delusions) was used to separate out a severely delusional schizophrenia group ( $n = 10$ ) from the mild/no delusions schizophrenia group ( $n = 40$ ) (delusions item = 4 or < 4 respectively). This was the only variable on which the two schizophrenia groups differed, with  $t$ -tests indicating no between group differences for age of onset of schizophrenia, length of illness, age first hospitalised or number of hospitalisations. The bipolar group rated low for delusions with 42 participants scoring a zero, 13

a one, and five a two for the SSPI delusions item. Patient participants were excluded if they had a current or past history of severe drug abuse or alcoholism, and if they had experienced a head injury or a concussion resulting in a loss of consciousness for 10 minutes or more. With the exception of one individual, all participants diagnosed with schizophrenia were stabilized with neuroleptic medications, with the large majority taking atypical neuroleptics. Of the patients in the psychiatric control group, 21 of 55 were taking atypical neuroleptics, while the remainder were not taking any antipsychotic medication. Chlorpromazine equivalents are presented in Table 1.

Forty-four healthy controls were recruited through advertisement and word-of-mouth. Screening with a medical questionnaire ensured that none of the healthy participants had any current or prior history of psychiatric illness. Additional exclusion criteria were the same as those employed for the patient groups.

All participants were fluent in English, used English daily, and had been speaking English daily for at least the past five years. Prior to participation in the BADE task, intelligence estimates were made using the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990) for current verbal and non-verbal intelligence, and the Ammons Quick Test (QUICK; Ammons & Ammons, 1962) for an assessment of premorbid IQ. The K-BIT takes approximately 20 minutes to administer and includes a verbal and non-verbal intelligence scale, as well as providing a composite IQ score. The verbal scale includes word knowledge, verbal concept formation, and reasoning ability, while the non-verbal scale includes a matrices subtest. The QUICK takes approximately five minutes to administer and requires participants to indicate which of four pictures best fits the meaning of a series of words spoken aloud by the experimenter. Socio-economic status was estimated using the Hollingshead Two-Factor Index of

Social Position (Hollingshead, 1957) using highest parental occupation and education level.

Sociodemographic characteristics of the sample are summarized in Table 1, and any significant between-groups differences on these variables are flagged. Where significant between-groups differences were found, those variables were covaried out of the main effects of interest using analysis of covariance (ANCOVA).

### *Procedure*

The BADE test was conducted in a testing room at the University of British Columbia (Canada) using Microsoft Internet Explorer 6.0. An experimenter was present throughout the entire procedure. The BADE test used 30 delusion-neutral scenarios (Buchy, Woodward & Liotti, 2007; Woodward, Buchy, Moritz & Liotti, 2007). Twenty-four of the BADE scenarios were expressed in the form of three sequentially presented and increasingly disambiguating sentences describing a delusion-neutral scenario. For example, one of the scenarios described a girl who can't fall asleep on Christmas Eve because, as becomes increasingly apparent, she is eagerly anticipating Santa Claus bringing Christmas presents later that night. The first sentence presented for this scenario was, "Jenny can't fall asleep". For this scenario, the four interpretations that must be rated for plausibility were: one "true" interpretation (e.g., "Jenny is excited about Christmas morning"), two "lure" interpretations (e.g. "Jenny is nervous about her exam the next day" and "Jenny is worried about her ill mother"), and one absurd interpretation (e.g., "Jenny loves her bed"). Note that the "true" interpretation would not be considered particularly plausible after the first sentence describing the scenario. The second sentence presented was, "Jenny can't wait until it is finally morning", and the third was, "Jenny wonders how many presents she will find under the tree". After the third sentence has been presented, the

scenario should be clear to the participants, indicating that the lures should be down-rated, and the true interpretation up-rated. Thus, the true interpretation appeared to be less plausible initially, though became increasingly plausible with each successive piece of confirmatory evidence, while the two lures initially appeared to be plausible, but became increasingly less plausible. The absurd interpretation was designed to be implausible even from the presentation of the initial sentence. This series of events is depicted in screen shots in Appendix I, with further BADE trial examples given in Appendix II. Plausibility ratings were recorded on a 0-10 rating scale using a scroll bar positioned beneath each interpretation. The nominal ratings “Poor”, “Possible”, “Good”, and “Excellent” were evenly distributed along the scale. Ratings were made and adjusted via mouse click. To avoid participants settling into one pattern of responding, the remaining six of the 30 scenarios were “distracters”. For these stimuli the true answer was apparent following the very first piece of evidence. The sentences for this task were chosen and constructed based on extensive piloting, and were previously rated by 69 undergraduate students (Buchy, Woodward & Liotti, 2007; Woodward, Buchy, Moritz & Liotti, 2007), confirming the desired lure/absurd/true pattern.

Prior to testing, the experimenter described the rating scale to participants, indicating that each of the interpretations were independent, and that ratings could be changed after the presentation of each piece of information. One practice trial was given to familiarize participants with the procedure.

### *Data Analysis*

In order to consider overlap and redundancy between test items to determine the number of cognitive operations underlying all responses, we submitted all 12 plausibility ratings (four

rating types - one true, two lure and one absurd interpretation - each rated three times after three presentations of evidence) to a principal component analysis (PCA). We planned a Group (schizophrenia, severely delusional schizophrenia, bipolar disorder, and healthy controls) by Component (a two-component solution was expected) to a mixed-model ANOVA. Significant main or interaction effects were followed up using univariate ANOVAs and *t*-tests.

## **Results**

### *Principal Component Analysis (PCA)*

The BADE test variables (mean plausibility ratings for true interpretations, absurd interpretations and both lures following each of the three pieces of evidence; Table 2) for all participants were entered into a PCA with oblimin rotation. The scree plot and Kaiser-Guttman criterion (eigenvalues > 1) converged on a clear two component solution (eigenvalues: 7.06, 3.10, 0.72, 0.46, 0.24, 0.13, 0.11, 0.08, 0.04, 0.03, 0.02 and 0.02), which accounted for 85% of the total variance. Component pattern weights for the total sample are listed in Table 3 (the same analysis carried out on each subsample is also listed in order to demonstrate the reliability of the solution). The rotation sums of squares loadings showed that Component 1 accounted for 46.96% of the total variance, and that the rotated Component 2 accounted for approximately 46.03% of the total variance (these percentages are over-additive due to the non-orthogonal rotation). This suggests that all BADE ratings can generally be interpreted within the context of two cognitive operations. Component 1 was dominated by plausibility ratings for all absurd interpretations, lure interpretations following sentence two and, in particular, sentence three, and true interpretations on rating three (loaded negatively). Component 2 was dominated by ratings

for lure and true interpretations after receiving the first and second pieces of evidence, as well as the true interpretation on rating three (loaded positively).

An important factor for labeling and distinguishing between the components is the true interpretation, rating three, which loads strongly on both components, but in opposite directions. For Component 1, these negative loadings indicate that people rating high for lures rate low for trues, and vice-versa, indexing the degree to which they have fully integrated the disambiguating evidence; thus, this component was labeled “Evidence Integration”. The justification for Component 2 is less clear from the items loading onto it. Investigation of scatter plots of pairs of the rating items that load high on this component shows that some people rate low when other people are rating high, causing these items loading on the component to cluster together due to correspondences between those individual differences. Thus, as a broad rule, all ratings that tended to be relatively high loaded onto this component, as correspondences between individual differences for generally low raters drove this component, and these correspondences between individual differences were lost when *all* participants were rating low. Thus, Component 2 was interpreted as “Conservatism”. Split-half reliability computed on the components described above (24 BADE items randomly split into 2 sets of 12) was exceptionally high, with a Pearson’s correlation coefficient of 0.94 for Conservatism and 0.93 for Evidence Integration. The two components were correlated at 0.25.

#### *Group by Component Analysis*

The component scores for Evidence Integration and Conservatism were entered into a repeated measures ANOVA to determine whether they were able to discriminate the severely delusional schizophrenia group from the other groups. This analysis revealed a significant

interaction of Group membership with Component,  $F(3,149) = 3.77, p < 0.05, \eta^2 = 0.07$ , reflecting a significant Group effect for Evidence Integration,  $F(3,149) = 8.24, p < 0.001, \eta^2 = 0.14$ , but no significant Group effect for Conservatism  $F(3,149) = 0.21, p = 0.88$ . Between groups  $t$ -tests for Evidence Integration indicated that the severely delusional schizophrenia group ( $M = 1.20$ ) had significantly higher component scores than the other schizophrenia group ( $M = 0.12$ ),  $t(47) = 2.58, p < 0.05, \eta^2 = 0.12$ , the psychiatric control group ( $M = 0.00$ ),  $t(68) = 3.23, p < 0.01, \eta^2 = 0.13$ , and the healthy control group ( $M = -0.39$ ),  $t(52) = 5.07, p < 0.001, \eta^2 = 0.33$ . Both the schizophrenia group,  $t(81) = 2.94, p < 0.01, \eta^2 = 0.1$ , and the psychiatric control group,  $t(102) = 2.49, p < 0.05, \eta^2 = 0.06$ , had significantly higher Evidence Integration component scores than the healthy control group, though were not significantly different from each other. These relationships are presented in Figure 1. ANCOVAs for demographic items for which significant group differences were found (Table 1) showed that the between groups differences on Evidence Integration remained significant when these variables were covaried out (Age,  $F(3,149) = 7.76, p < 0.001, \eta^2 = 0.13$ ; K-BIT Vocabulary,  $F(3,149) = 6.04, p = 0.001, \eta^2 = 0.11$ ; K-BIT Comprehensive,  $F(3,149) = 6.44, p < 0.001, \eta^2 = 0.11$ ; Quick Test,  $F(3,149) = 8.95, p < 0.001, \eta^2 = 0.15$ ; Education,  $F(3,149) = 8.46, p < 0.001, \eta^2 = 0.15$ ). The specificity of the Evidence Integration component as a discriminator between the schizophrenia group and severely delusional schizophrenia group was examined by repeating the analysis described above for high and low hallucinations, and high and low formal thought disorder, using comparable subdivisions of the same schizophrenia patient sample. Neither Evidence Integration nor Conservatism differed between high and low subgroups for hallucinations (all  $ps > 0.13$ ) or formal thought disorder (all  $ps > 0.12$ ).

## **Discussion**

Two components accounted for 85% of the variance in the BADE task: Evidence Integration (plausibility ratings reflecting the degree to which disambiguating information has been integrated) and Conservatism (plausibility ratings indicating a reduced willingness to provide high ratings when justified). Only Evidence Integration scores differed between the severely delusional schizophrenia group and the other groups; Conservatism scores did not differ, arguing against a generalized performance deficit. Delusions were associated with higher Evidence Integration scores, reflecting comparatively higher plausibility ratings for disconfirmed interpretations, combined with lower ratings for confirmed (true) interpretations, reflecting a reduced tendency to adjust beliefs when confronted with disambiguating evidence. The opposite was true for the healthy control group, who showed more complete evidence integration than any of the patient groups.

The two components derived from the PCA of our previous BADE study on healthy subjects agreed with those reported here (Woodward, Moritz, Menon & Klinge, 2008), but as only lure items were analyzed in that study, the selection of component labels did not benefit from observation of how true and absurd items load onto the components along with the lures. In our past study, we labeled Component 1 “Integration of Disconfirmatory Evidence”, but due the high involvement of the absurd items, and the negative loadings on the true items, we changed the name to Evidence Integration (that is, integration of disambiguating evidence, disconfirmatory or otherwise). In our past study, we labeled Component 2 “Initial Belief”, but due the high involvement of the true items on ratings two and three, and the evidence that this component is caused by participants rating a number of items low when there is reason to rate high, we renamed this component “Conservatism”.

The current results show support for the BADE effect on the lure items (i.e., an inadequate downward adjustment of plausibility ratings when necessitated by incoming disambiguating evidence), but suggest that the BADE effect is caused by a more general effect that also affects ratings of the absurd items, and reciprocally, the true items. The cognitive operation that we propose to be underlying the Evidence Integration component, and the BADE aspect of this component, is “hypersalience of evidence-hypothesis matches” (Speechley, Whitman & Woodward, 2010), which holds that when the encountered evidence matches a hypothesis currently held in mind, people with schizophrenia and severe delusions give inordinately high weight to that evidence. The hypersalience of evidence-hypothesis matches theory was derived from a previous study where we observed that a severely delusional schizophrenia group showed a tendency to overrate any option supported by the current evidence. In the BADE task, participants were instructed to rate multiple interpretations independently, such that all interpretations could be considered correct simultaneously. Under these circumstances, where “there is no wrong answer”, hypersalience of evidence-hypothesis matches may lead to an enhancement of weak matches (i.e., absurd and disconfirmed interpretations) due to giving inordinately high weight to weak evidence that is perceived to match the available interpretation, leading to an unwillingness to fully integrate disambiguating information. In this way, the Evidence Integration component of the BADE task may partially reflect a tendency to give too much consideration to disconfirmed or implausible explanations owing to hypersalience of evidence-hypothesis matches.

Hypersalience of evidence-hypothesis matches is seen to contribute to the genesis of delusions due to overvaluation (hypersalience) of matches between a delusional idea (e.g., I think the CIA is spying on me) and evidence (e.g., people are staring at me) at the expense of

disconfirming evidence. This can lead to delusions that cannot be properly disconfirmed (e.g., that fact that suspected recording devices are not found in the heat vents can be ignored, because crackling on the phone line is a highly salient confirmation of the CIA delusion). Delusions may be elaborated as evidence-hypothesis matches broaden and intensify (e.g., a black van passing by slowly suggests it is driven by an observant CIA agent). The delusions cannot be properly disconfirmed because the balance between confirming and disconfirming evidence is skewed away from disconfirming evidence. Trait impairments relating to executive functions, memory, attention, and poor functional outcome may exacerbate these processes (Green, Kern, Braff & Mintz, 2000).

This study is subject to a number of limitations. First, the sample size for the severely delusional schizophrenia group is small. However, the ability to detect the effects reported here depended upon procuring a relatively large sample of people diagnosed with schizophrenia ( $N = 50$ ), such that a reasonably large sample of patients with the most severe delusion symptoms could be split off from the rest of the group. In all of our past BADE studies, total sample sizes for the schizophrenia group were smaller, requiring lower cut-offs for inclusion in the delusion group in order to allow reasonable subgroup sample sizes. Lower cut offs resulted in the averaging together of patients experiencing milder levels of delusions with those experiencing severe delusions, perhaps accounting for some of the inconsistency previously reported when investigating whether a BADE effect is more pronounced in a delusional schizophrenia group compared to a non-delusional group. Another potential limitation of this study is that there were a number of differences observed between the groups on demographic variables. Despite using ANCOVA to ensure that these differences did not impact the main results, ideally all groups would be matched on all demographic variables. However, in this regard it is very important to

note that the key group comparison, the schizophrenia versus severely delusional schizophrenia group, was not affected by this limitation, as these two groups did not differ significantly on any demographic variables.

The results of this study support the notion that the BADE task may engage two distinct cognitive processes, though only the Evidence Integration component, which requires the consideration, or reconsideration, of beliefs in light of disconfirmatory evidence, appears to discriminate between the severely delusional schizophrenia group and other psychiatric patient groups. Evidence integration and the BADE have been incorporated as treatment targets in metacognitive training (MCT; Moritz & Woodward, 2010). MCT is designed to provide patients with knowledge of their metacognitive infrastructure, for example, reasoning biases such as BADE, and provides exercises and experience in correcting these biases. Increased awareness of the role of Evidence Integration in reasoning may help protect against delusional ideation in schizophrenia. Ultimately, this study provides further confirmation that the cognitive biases in delusions extend beyond material congruent with an individual's specific delusions to neutral, unrelated content, and as such, may reflect a pervasive reasoning deficit predisposing towards the formation and maintenance of delusional ideation.

## **Acknowledgments**

The authors thank Ligaya Allmer, Jessica Bristowe, Lisa Buchy, Tanya Fawkes, Sara Flann, Nathalie Hovatta, Tonya Kragelj, Rachel McKay, Izabella Patyk, and Rachel Richards for assistance with data collection, task design and data management.

## References

- Ammons, R.B. & Ammons, C.H.** (1962). The quick test (QT) provisional manual. *Psychological Reports*, **11**, 111-161.
- Barrett, S.L., Mulholland, C.C., Cooper, S.J. & Rushe, T.M.** (2009). Patterns of neurocognitive impairment in first-episode bipolar disorder and schizophrenia. *British Journal of Psychiatry*, **195**, 67-72.
- Buchy, L., Woodward, T.S. & Liotti, M.** (2007). A cognitive bias against disconfirmatory evidence (BADE) is associated with schizotypal traits. *Schizophrenia Research*, **90**, 334–337.
- Green, M.F., Kern, R.S., Braff, D.L., Mintz, J.** (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "Right stuff"? *Schizophrenia Bulletin*, **26**, 119-136.
- Hollingshead, A.** (1957). *Two-factor index of social position*. Yale University Press.
- Kaufman, A.S. & Kaufman, N.L.** (1990). *Kaufman Brief Intelligence Test Manual*. American Guidance Service, Circle Pines, MN.
- Liddle, P.F., Ngan, E.T.C., 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.
- Moritz, S. & Woodward, T.S.** (2006). A generalized bias against disconfirmatory evidence in schizophrenia. *Psychiatry Research*, **142**, 157-65.
- Moritz, S. & Woodward, T.S.** (2007). Metacognitive training in schizophrenia: from basic research to knowledge translation and intervention. *Current Opinion in Psychiatry*, **20**, 619–625.

**Moritz, S. & Woodward, T.S.** (2010). *Metacognitive training for patients with schizophrenia (MCT). Manual*, 2nd edn. VanHam Campus Verlag: Hamburg.

**Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Barker, R. & Dunbar, G.C.** (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview. *Journal of Clinical Psychiatry*, **59**, 22-33.

**Speechley, W.J., Whitman, J.C. & Woodward, T.S.** (2010). The contribution of hypersalience to the “jumping to conclusions” bias associated with delusions in schizophrenia. *Journal of Psychiatry and Neuroscience*, **35**, 7-17.

**Woodward, T.S., Moritz, S. & Chen, E.Y.H.** (2006). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions: a study in an Asian sample with first episode schizophrenia spectrum disorders. *Schizophrenia Research*, **83**, 297–298.

**Woodward, T.S., Moritz, S., Cuttler, C. & Whitman, J.** (2006). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions in schizophrenia. *Journal of Clinical and Experimental Neuropsychology*, **28**, 605–617.

**Woodward, T.S., Moritz, S., Menon, M. & Klinge, R.** (2008). Belief inflexibility in schizophrenia. *Cognitive Neuropsychiatry*, **13**, 267-277.

**Woodward, T.S., Buchy, L., Moritz, S. & Liotti, M.** (2007). A bias against disconfirmatory evidence is associated with delusion proneness in a nonclinical sample. *Schizophrenia Bulletin*, **33**, 1023-1028.

**Table 1.** Sociodemographic characteristics of the samples: means, with standard deviations in parentheses.

	Healthy controls ( <i>n</i> = 44)	Bipolar ( <i>n</i> = 60)	Non-delusional sz ( <i>n</i> = 40)	Delusional sz ( <i>n</i> = 10)
Age	35.39 (11.37)	41.62 (10.05) <sup>††¶¶</sup>	35.36 (11.78)	41.10 (12.19)
Sex (M:F)	17:27 <sup>‡</sup>	27:33 <sup>††</sup>	30:10	5:5
Years of Education	14.52 (2.21) <sup>§††</sup>	14.90 (3.48) <sup>*†††</sup>	12.33 (2.86)	12.00 (1.89)
Length of illness	n/a	14.64 (11.07)	13.00 (10.87)	20.33 (10.36)
K-BIT vocab	100.48 (15.66)	105.47 (17.00) <sup>*††</sup>	93.97 (11.81)	90.33 (12.03)
K-BIT matrices	98.88 (22.37)	97.44 (12.69)	94.86 (14.33)	94.00 (19.07)
K-BIT composite	99.73 (17.69)	100.86 (11.53) <sup>*††</sup>	93.97 (12.63)	91.67 (16.85)
QUICK	102.60 (11.68)	108.42 (11.33) <sup>*†¶¶</sup>	103.42 (11.58)	98.67 (8.19)
Social status	35.74 (16.44)	31.21 (17.35)	33.71 (16.09)	42.33 (19.59)
Chlorpromazine Equivalent (mg)	n/a	110.05 (479.27)	230.59 (352.29)	301.88 (186.51)

\*Delusional schizophrenia vs. bipolar,  $p < 0.05$ ; §Delusional schizophrenia vs. healthy,  $p < 0.01$ ;

†Non-delusional schizophrenia vs. bipolar,  $p < 0.05$ ; †† Non-delusional schizophrenia vs. bipolar,

$p < 0.01$ ; ††† Non-delusional schizophrenia vs. bipolar,  $p < 0.001$ ; ‡Non-delusional schizophrenia

vs. healthy,  $p < 0.01$ ; ‡‡Non-delusional schizophrenia vs. healthy,  $p < 0.001$ ; ¶Bipolar vs. healthy,

$p < 0.05$ ; ¶¶Bipolar vs. healthy,  $p < 0.01$

**Table 2.** Mean plausibility ratings for all BADE test variables following each of the three pieces of evidence; means, with standard deviations in parentheses

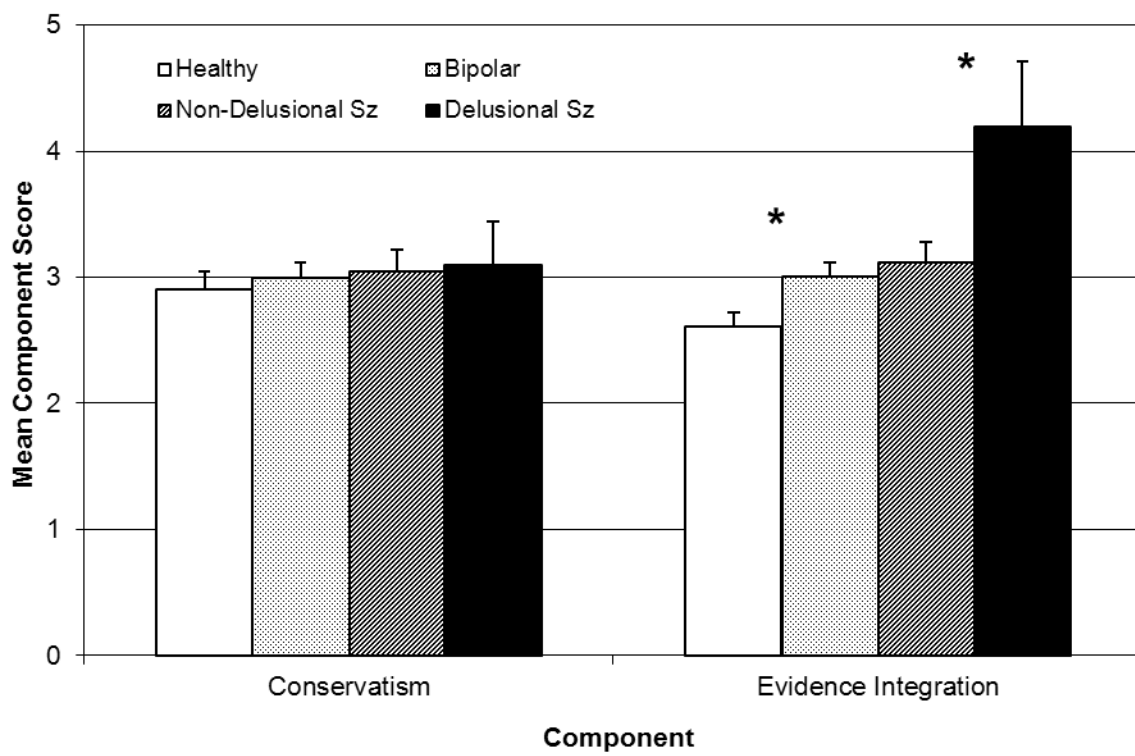
Statement	Delusional Sz	Schizophrenia	Bipolar	Healthy Control
Absurd One	2.59 (2.32)	1.65 (1.30)	1.67 (1.30)	1.05 (0.82)
Absurd Two	2.83 (2.34)	1.41 (1.15)	1.35 (1.19)	0.71 (0.66)
Absurd Three	2.44 (2.33)	0.88 (1.10)	0.94 (1.09)	0.33 (0.49)
Lure A One	5.45 (2.12)	5.14 (2.15)	4.95 (2.05)	4.60 (2.05)
Lure A Two	4.66 (2.24)	3.96 (1.79)	3.41 (1.50)	3.34 (1.58)
Lure A Three	3.15 (2.65)	1.97 (1.98)	1.45 (1.23)	1.15 (1.28)
Lure B One	5.34 (2.24)	4.67 (2.14)	4.66 (2.03)	4.22 (1.96)
Lure B Two	4.67 (2.27)	3.70 (1.73)	3.56 (1.63)	3.28 (1.43)
Lure B Three	3.46 (2.59)	1.85 (1.82)	1.63 (1.46)	1.14 (1.29)
True One	5.19 (2.47)	4.82 (2.12)	4.71 (2.01)	4.34 (1.94)
True Two	6.48 (2.49)	6.54 (2.12)	6.58 (1.93)	6.36 (1.74)
True Three	8.24 (2.34)	9.07 (1.36)	9.09 (1.55)	9.40 (0.68)

**Table 3.** Pattern weights for BADE variables ( $N = 149$ ). The participants were comprised of a schizophrenia group ( $n = 50$ ), a psychiatric control group ( $n = 60$ ) and a healthy control group ( $n = 44$ ).

Statement	Component							
	1				2			
	Evidence Integration				Conservatism			
	T	S	B	C	T	S	B	C
Absurd one	<b>.775</b>	<b>.796</b>	<b>.798</b>	<b>.532</b>	.291	.264	.300	<b>.446</b>
Absurd two	<b>.908</b>	<b>.900</b>	<b>.926</b>	<b>.835</b>	.098	.114	.097	.090
Absurd three	<b>.961</b>	<b>.977</b>	<b>.958</b>	<b>.911</b>	-.116	-.135	-.098	-.245
Lure A one	.185	.278	.123	.075	<b>.895</b>	<b>.835</b>	<b>.928</b>	<b>.941</b>
Lure A two	<b>.552</b>	<b>.656</b>	<b>.529</b>	<b>.441</b>	<b>.641</b>	<b>.541</b>	<b>.692</b>	<b>.716</b>
Lure A three	<b>.901</b>	<b>.952</b>	<b>.868</b>	<b>.875</b>	.014	-.117	.094	.023
Lure B one	.088	.165	.008	.037	<b>.939</b>	<b>.890</b>	<b>.973</b>	<b>.963</b>
Lure B two	<b>.533</b>	<b>.626</b>	<b>.495</b>	<b>.420</b>	<b>.628</b>	<b>.543</b>	<b>.653</b>	<b>.724</b>
Lure B three	<b>.882</b>	<b>.925</b>	<b>.878</b>	<b>.860</b>	.009	-.066	-.023	.059
True one	.088	.130	.041	.045	<b>.942</b>	<b>.920</b>	<b>.960</b>	<b>.952</b>
True two	-.141	-.071	-.248	-.059	<b>.965</b>	<b>.963</b>	<b>.957</b>	<b>.946</b>
True three	<b>-.584</b>	<b>-.462</b>	<b>-.684</b>	<b>-.426</b>	<b>.645</b>	<b>.729</b>	<b>.553</b>	<b>.758</b>

*Note:* All pattern weights over 0.40 are set in bold font. T = total sample; S = schizophrenia group; B = bipolar group; C = healthy control group

**Figure 1.** Mean component scores (Conservatism and Evidence Integration) plotted as a function of group (healthy controls ( $n = 44$ ), psychiatric controls ( $n = 60$ ), non-delusional schizophrenia patients ( $n = 40$ ) and delusional schizophrenia patients ( $n = 10$ )). The mean of the component scores was rescaled to 3 for display purposes.



\*  $p < 0.05$  for adjacent bars.

## Appendix I: BADE task screen shots from a single trial.

Screen 1 of 3

Jenny can't fall asleep.

---

Jenny is excited about Christmas morning.

←  →

0 1 2 3 4 5 6 7 8 9 10  
Poor Possible Good Excellent

Jenny is worried about her ill mother.

←  →

0 1 2 3 4 5 6 7 8 9 10  
Poor Possible Good Excellent

Jenny loves her bed.

←  →

0 1 2 3 4 5 6 7 8 9 10  
Poor Possible Good Excellent

Jenny is nervous about her exam the next day.

←  →

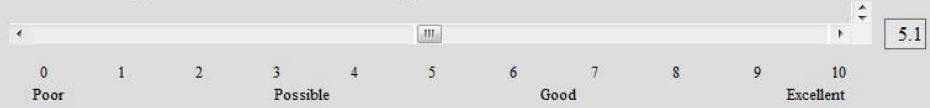
0 1 2 3 4 5 6 7 8 9 10  
Poor Possible Good Excellent

Screen 2 of 3

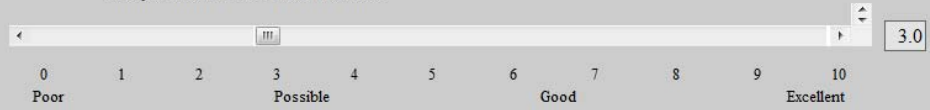
Jenny can't fall asleep.

Jenny can't wait until it is finally morning.

Jenny is excited about Christmas morning.



Jenny is worried about her ill mother.



Jenny loves her bed.



Jenny is nervous about her exam the next day.



Screen 3 of 3

Jenny can't fall asleep.

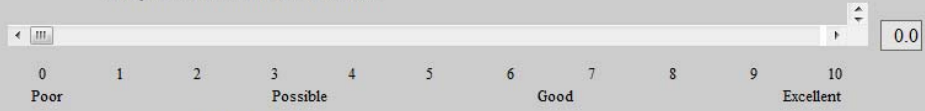
Jenny can't wait until it is finally morning.

Jenny wonders how many presents she will find under the tree.

Jenny is excited about Christmas morning.



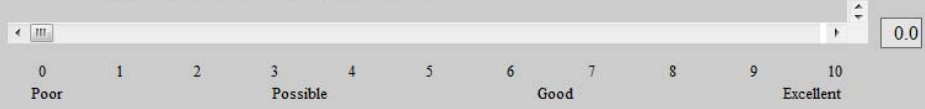
Jenny is worried about her ill mother.



Jenny loves her bed.



Jenny is nervous about her exam the next day.



## Appendix II: Examples of BADE trials.

Example	Scenario Statements	Interpretations
1	Nicholas is driving his car very fast Nicholas did not stop at the red light Nicholas injured a little girl with his car	Nicholas is running late for work (L) Nicholas' wife is in labour (L) Nicholas hates going for walks (A) Nicholas is a hit and run offender (T)
2	Amanda is very thin Amanda has a difficult life Amanda doesn't even have a home	Amanda is a runway model (L) Amanda has an eating disorder (L) Amanda has lost her fake teeth (A) Amanda is homeless (T)
3	The woman has been in severe pain all day The woman is impatiently waiting for her special day to come The woman has a big belly	The woman is training to be a champion gymnast (L) The woman has only days left to live (L)  The woman loves to be tickled (A) The woman is about to have a baby (T)
4	Mark often comes home late from work Mark tells his wife that he is not in the mood to spend time with her Mark's wife is suspicious	Mark is trying to earn a promotion at work (L) Mark must work hard in order to not lose his job (L)  Mark thinks that money grows on trees (A) Mark is having an affair (T)