Delusion-Like Beliefs and Data Quality: Are Classic Cognitive Biases Artifacts of Carelessness?

Justin Sulik1*, Robert M. Ross2*, Ryan Balzan3, and Ryan McKay4

1 Cognition, Values & Behavior, Ludwig Maximilian University of Munich, Germany
2 Department of Psychology, Macquarie University, Australia
3 College of Education, Psychology and Social Work, Flinders University, Australia
4 Department of Psychology, Royal Holloway, University of London, UK

* these authors contributed equally to this work

Author Note

Justin Sulik https://orcid.org/0000-0002-0978-9496

Robert M. Ross https://orcid.org/0000-0001-8711-1675

Ryan Balzan https://orcid.org/0000-0003-2007-4317

Ryan McKay https://orcid.org/0000-0001-7781-1539

We have no conflicts of interest to disclose. The study passed the internal ethics procedure at Royal Holloway, University of London. A repository of data, materials, analyses and preregistrations is at https://osf.io/fvkcn. An earlier version of this manuscript was made publicly accessible on 26th April 2021 on the PsyArXiv pre-print server: https://psyarxiv.com/ntsve/. We thank Eliane Deschrijver for helpful suggestions on earlier drafts. Correspondence concerning this article should be addressed to Justin Sulik, Cognition Values & Behavior LMU, Gabelsbergerstr 26, Munich 81333, Germany. Email: justin.sulik@gmail.com.
Abstract

There is widespread agreement that delusions in clinical populations and delusion-like beliefs in the general population are, in part, caused by cognitive biases. Much of the evidence comes from two influential tasks: the Beads Task and the Bias Against Disconfirmatory Evidence (BADE) Task. However, research using these tasks has been hampered by conceptual and empirical inconsistencies. In an online study, we examined relationships between delusion-like beliefs in the general population and cognitive biases associated with these tasks. Our study had four key strengths: a new animated Beads Task designed to reduce task miscomprehension, several data-quality checks to identify careless responders, a large sample (n = 1,002), and a pre-registered analysis plan. When analyzing the full sample, our results replicated classic relationships between cognitive biases and delusion-like beliefs. However, when we removed 82 careless participants from the analyses (8.2% of the sample) we found that many of these relationships were severely diminished and, in some cases, eliminated outright. These results suggest that some (but not all) seemingly well-established relationships between cognitive biases and delusion-like beliefs might be artifacts of careless responding.

Keywords: Beads Task; Bias Against Disconfirmatory Evidence (BADE); cognitive bias; continuum model; delusion; jumping to conclusions (JTC)
General Scientific Summary (GSS)

Research suggests that cognitive biases play a key role in the development of delusion-like beliefs. For instance, participants who endorse such beliefs have been reported to “jump to conclusions” when performing abstract data-gathering tasks and to display a “bias against disconfirmatory evidence” when determining the best explanation for a scenario. However, the present study suggests that some (but not all) seemingly well-established relationships between cognitive biases and delusion-like beliefs might, in fact, be spurious—driven by careless responding in a subset of research participants.
Delusion-Like Beliefs and Data Quality: Are Classic Cognitive Biases Artifacts of Carelessness?

Research points to the existence of a psychosis continuum such that delusions in patients with psychotic disorders appear in less severe forms as delusion-like beliefs in the general population (Linscott & van Os, 2013; van Os et al., 2009). A large body of evidence suggests that cognitive biases are implicated in the development and maintenance of delusions and delusion-like beliefs (Dudley et al., 2016; McLean et al., 2017; Ross et al., 2015; So et al., 2016). However, there is reason to be cautious when interpreting this evidence. Notably, much of the evidence for an association between cognitive biases, delusions and delusion-like beliefs comes from research using two influential tasks—the Beads Task and the Bias Against Disconfirmatory Evidence (BADE) Task—but, as we shall argue, this research suffers from a range of conceptual and empirical issues.

In the Beads Task (Huq et al., 1988), participants are shown two urns filled with beads in complementary colors: for instance, one labelled “mainly pink” and the other labelled “mainly green”. The experimenter hides the urns and ostensibly draws beads from one of them. After each draw, participants are asked if they are ready to guess which urn the beads are being drawn from or if they would like to see another bead. Three cognitive biases have been identified using this task. First, delusions and delusion-like beliefs are associated with requesting fewer beads before making a decision about which urn the beads are being drawn from, suggesting a “jumping to conclusions” (JTC) bias (Dudley et al., 2016; McLean et al., 2017; Ross et al., 2015; So et al., 2016). Second, when participants are asked to rate how likely it is that the beads are coming from each urn after each successive draw, delusions and delusion-like beliefs are associated with “overadjusting” to a newly revealed bead that contradicts the previously favored hypothesis, suggesting an “overadjustment bias” (Garety et
al., 1991; Rodier et al., 2011). Third, when participants are asked to rate their confidence in their decision about which urn the beads are being drawn from once they have made the decision to stop drawing beads, delusions and delusion-like beliefs are associated with relatively low subjective confidence, suggesting a “liberal acceptance bias” (Moritz et al., 2006; Moritz et al., 2007).

In the BADE Task (Woodward et al., 2006), participants read four competing explanations of a scenario. At each stage of the task, they are provided with additional information about the scenario and are asked to rate the plausibility of the competing explanations: two “lure” explanations which initially seem plausible but become less plausible as more information is revealed; a “true” explanation which becomes the most plausible as more information is revealed; and an “absurd” explanation which is consistently implausible. Two putative cognitive biases are routinely studied using this task. First, delusions and delusion-like beliefs have been associated with favoring the initially plausible lure explanations even as evidence against them mounts, suggesting a “bias against disconfirmatory evidence” (BADE; McLean et al., 2017; Menon et al., 2013). Second, delusions and delusion-like beliefs have been associated with overrating the plausibility of the absurd explanation, also termed a “liberal acceptance bias” (Woodward et al., 2006).

A concern with this literature is that appeals to cognitive biases as drivers of delusions and delusion-like beliefs lack explanatory depth. For instance, it has been argued that “portrayal of people with delusions as having a JTC bias is a redescription rather than an explanation” (Corlett & Fletcher, 2014, p. 399). Analogous criticisms can be applied to the other aforementioned cognitive biases. For example, the hypothesis that people with delusions display a bias against disconfirmatory evidence seems true almost by definition—the American Psychiatric Association defines delusions as “fixed beliefs that are not
amenable to change in light of conflicting evidence” (APA, 2022, p. 101). One potential path forward is to situate proposed cognitive biases within a more general cognitive model of information processing, as a number of researchers have done (Bronstein, Pennycook, Joormann, et al., 2019; Coltheart & Davies, 2021; Connors & Halligan, 2020).

Dual process theories of thinking are influential in cognitive science. These theories distinguish two broad types of thinking (Evans & Stanovich, 2013): Type 1 or “intuitive” processing, which is relatively fast and effortless, and which has been characterized as a “machine for jumping to conclusions” (Kahneman, 2011, p. 79); and Type 2 or “analytic” processing, which is relatively slow and effortful. It has been proposed that dual process theories provide a useful cognitive model for understanding delusions (Bronstein, Pennycook, Joormann, et al., 2019; Ross et al., 2016). In particular, people with delusions or delusion-like beliefs may fail to critically engage effortful Type 2 analytic processing to evaluate the outputs of Type 1 intuitive processing. Supporting this proposal, poor performance on a test of analytic thinking has been associated with schizophrenia in clinical research (Puveendrakumaran et al., 2020); and with delusion-like beliefs (Bronstein, Pennycook, Bear, et al., 2019), the JTC bias (Ross et al., 2016), and the BADE (Bronstein, Everaert, et al., 2019) in general population studies.

In the present study, we had participants complete the Beads Task, the BADE Task, a test of analytic thinking, and a measure of delusion-like beliefs in order to rigorously examine relationships among the underlying constructs. Specifically, we aimed to address four key inconsistencies in the empirical literature.

First, there is mixed empirical evidence concerning the proposed integration between delusion-like beliefs and dual process theories of thinking. Although one large study found a relationship between delusion-like beliefs and analytic thinking (Bronstein, Pennycook, Bear,
et al., 2019), another large study did not (Ross et al., 2016). A further study found that analytic thinking predicted performance in the BADE Task in the general population (Bronstein, Everaert, et al., 2019), but that study did not measure delusion-like beliefs. As dual process theories aim to add explanatory depth to research on cognitive biases, we would expect that analytic thinking predicts the same biases that, in turn, predict delusion-like beliefs. Hence, Research Question 1: Does analytic thinking predict the JTC bias and the BADE and, in turn, do these biases predict delusion-like beliefs?¹

Second, there is a striking tension between cognitive biases identified using the Beads Task and the BADE Task (Furl et al., in press). As noted above, in the Beads Task, delusions and delusion-like beliefs are associated with a tendency to over-accommodate evidence that disconfirms a pre-established pattern (i.e., the overadjustment bias; Garety et al., 1991; Rodier et al., 2011). By contrast, in the BADE Task, delusions and delusion-like beliefs are associated with a tendency to under-accommodate disconfirmatory evidence (i.e., the BADE; McLean et al., 2017). It seems incongruous that delusions and delusion-like beliefs should be associated with over-accommodating evidence in one task and under-accommodating evidence in another. However, given that miscomprehension of the Beads Task appears to be common (Balzan, Delfabbro, & Galletly, 2012; Balzan, Delfabbro, Galletly, et al., 2012; Howe et al., 2018), it is uncertain whether this apparent contradiction really exists. Hence, Research Question 2: Are the overadjustment bias and the BADE associated with each other and with delusion-like beliefs?

¹ This research question differs from that reported in our pre-registration. Our pre-registered research question was inspired by an analysis conducted by Ross et al. (2016) and asked whether analytic thinking predicts the JTC bias and the BADE while controlling for delusion-like beliefs. However, upon further reflection we realized that this analysis does not test a cognitive model. For this reason, we developed a new research question and analysis plan that focuses on the cognitive model developed by Bronstein, Pennycook, Joorman et al. (2019). This modified research question asks whether these cognitive biases mediate the relationship between analytic style and delusion-like beliefs. The original pre-registered analysis is reported in SM03.
Third, although the Beads Task and the BADE Task have occasionally been used together in a single study (Howe et al., 2018; Moritz et al., 2010; So et al., 2023), it is uncertain whether the JTC bias and the BADE jointly predict delusions or delusion-like beliefs. Hence, Research Question 3: Do the JTC bias and the BADE each predict unique variance in delusion-like beliefs, and if so, which is the stronger predictor?

Fourth, the liberal acceptance bias appears to be conceptualized in different ways in the Beads Task and the BADE Task. Liberal acceptance biasBEADS is operationalized as deciding to stop drawing beads while expressing low subjective confidence about which urn the beads are being drawn from (Moritz et al., 2016; Moritz et al., 2006; Moritz et al., 2007). By contrast, liberal acceptance biasBADE is operationalized as overrating the plausibility of an absurd interpretation of events (McLean et al., 2017). These two operationalizations are thought to measure the same underlying cognitive bias because they each appear to be driven by people readily accepting explanations on the basis of little evidence. Nonetheless, it is not clear whether these alternative operationalizations of a “liberal acceptance bias” really do measure the same underlying cognitive bias or if they are different biases that have been assigned the same name. Hence, Research Question 4: Are the liberal acceptance biasBEADS and the liberal acceptance biasBADE associated with each other and with delusion-like beliefs?

In examining these hypotheses, our study includes four notable strengths that are particularly relevant when considering meta-analyses of the Beads Task and the BADE Task literature (Dudley et al., 2016; McLean et al., 2017; Ross et al., 2015; So et al., 2016). First, to address concerns that misunderstanding of the Beads Task may impact results (Balzan,

---

2 For Hypothesis 4 we pre-registered an intention to test for an association between liberal acceptance biasBEADS and the liberal acceptance biasBADE, but a point of oversight was that we neglected to pre-register an intention to test for a relationship between these biases and delusion-like beliefs.
Delfabbro, & Galletly, 2012; Balzan, Delfabbro, Galletly, et al., 2012; Howe et al., 2018), we developed a new animated version of this task that specifically aims to facilitate comprehension and that includes a series of comprehension checks immediately after the demonstration. This is important because few—if any—of the Beads Task studies included in the aforementioned meta-analyses animated the task. A key advantage of our animated task is that we show participants the entire process of setting up the urns and drawing beads. Non-animated presentations of the Beads Task can be cognitively demanding and confusing because participants must read (or listen to) and understand complex descriptions without clear visual demonstrations.

Second, we included multiple data-quality checks throughout the study to examine whether “careless responding” might have biased results in this literature. We follow Arthur et al. (2021) in using the term “careless responding” to include “inattentiveness, arbitrary response patterns, or unwillingness to comply with the task demands” (p. 106). Data quality checks are important because careless responding can have a profound impact on results Arthur et al. (2021), including in clinical research (Agley et al., 2022; Chandler et al., 2020; Moseley et al., 2021; Ophir et al., 2020; Zorowitz et al., 2021).

Third, we recruited a large sample, enabling us to detect small effect sizes. This is important because meta-analyses are in agreement that some key effect sizes in this literature are small (Dudley et al., 2016; McLean et al., 2017; Ross et al., 2015; So et al., 2016).

Fourth, we pre-registered our analysis plan, share our raw data, and report our analysis scripts to make our research maximally transparent and to protect against incorrect inferences, which are issues of concern in the clinical literature (Tackett et al., 2019).

---

3 In fact, there is compelling evidence that surprisingly large numbers of people are willing to lie about their beliefs in surveys in certain contexts (Ross & Levy, 2023).
Methods

Pre-registrations, raw data, data processing scripts, analysis scripts, the full list of tasks with supplemental methods information, and an interactive study demonstration are available in Supplementary Materials SM01-SM11 on the Open Science Framework (OSF) project repository: https://osf.io/fvkcn/

Participants

We pre-registered a sample size of 1,000. A Bayes Factor calculation (see SM08) shows that \( n = 949 \) would be sufficient to provide positive evidence (BF \( \geq 3 \)) for a correlation \( r = .10 \), which was the mean correlation between JTC and delusion-like beliefs in a recent meta-analysis (Ross et al., 2015).

We recruited from Amazon’s Mechanical Turk service, a platform which is regularly used in clinical studies (Chandler & Shapiro, 2016). Participation was managed by Turkprime (Litman et al., 2017) which overshot by two (\( n = 1002 \)). Mean age was 35.2 years (SD = 10.8), including 507 males, 482 females, 3 non-binary people, and 10 who opted out of the gender question. The study was conducted in accordance with the APA ethical principles and passed the internal ethics procedure at Royal Holloway, University of London.

Procedure

After providing informed consent, participants completed the following key tasks.

Beads Task (Huq et al., 1988). Participants were presented with two urns full of beads: for example, one “majority pink” urn (containing 60 pink beads and 40 green beads) and one “majority green” urn (60 green beads; 40 pink beads). One urn was randomly selected, and participants were told that beads would be drawn from that single urn one at a time with replacement (in reality, the sequence was fixed). Before seeing any beads being drawn, and again after each draw, participants provided confidence ratings on a 101-point
scale ranging from 100% confidence in the “majority pink” urn at one end to 100%
confidence in the “majority green” urn at the other. We animated the task so that participants
could straightforwardly see that one urn was (ostensibly) picked at random and that beads
were drawn with replacement from that single urn. In addition, as per standard practice, to
minimize memory demands, previously drawn beads were displayed onscreen.\textsuperscript{4} Importantly,
participants only proceeded to the trials once they scored 100% on a four-item
comprehension test, otherwise the demonstration and test were repeated until this threshold
was reached. 85\% of participants passed all four Beads Task comprehension check questions
on their first attempt.

Participants completed two Beads Task trials. One was a “draws-to-decision” version
of the Beads Task (Huq et al., 1988), where participants could draw as many beads as they
wanted (up to a maximum of 30) before deciding the majority color of the urn. Draws-to-
decision was scored as the number of beads a participant drew, and the JTC bias was
conceptualized in terms of fewer draws to decision. Liberal acceptance bias\textsubscript{BEADS} was scored
as a participant’s confidence rating at the point they decided on the majority color (Moritz et
al., 2010), where lower confidence indicates more liberal acceptance. The other trial was a
“graded-estimates” version of the Beads Task (Garety et al., 1991) where participants always
drew ten beads, with beads of the minority color appearing on draws 3, 6 and 9. The
overadjustment bias was scored as the sum of the changes in a participant’s confidence each
time a bead of the minority color was drawn (a positive value indicating change towards the
minority color).

\textsuperscript{4} Our animated Beads Task is extremely flexible, which could facilitate re-use in future studies. By changing
simple parameters in the open-source code (available on the OSF project repository), the number of beads and
ratio of beads colors can be changed, the sequence of drawn beads can be specified or randomized, the record of
previous draws can be shown or hidden, feedback can be provided, and the response format can be changed.
Bias Against Disconfirmatory Evidence (BADE) Task (Woodward et al., 2006). Participants completed eight focal trials plus three filler trials (Buchy et al., 2007). In each trial, participants were shown four candidate explanations of a given scenario. Over three stages, more information was revealed about the scenario. After each reveal, participants rated how likely each explanation was, given the current information. Focal trials were designed to have two “lure” explanations which initially seemed plausible, but became less plausible as more information was revealed; a “true” explanation which became the most plausible as more information was revealed; and an “absurd” explanation which was consistently implausible. Here, we scored the BADE using the method of Sanford et al. (2014). For further details about this scoring method and discussion of alternative scoring methods developed by Bronstein and Cannon (2018), see SM03, SM05, SM06. Liberal acceptance bias was scored as the average of the ratings of absurd interpretations across all three stages of the task (Woodward et al., 2007) with a more positive score coded to indicate more bias (i.e., more readiness to accept implausible response options).

Peters et al. Delusions Inventory (PDI; Peters et al., 2004). Participants answered “yes” or “no” to 21 questions about delusion-like experiences, such as “Do you ever feel as if people seem to drop hints about you or say things with a double meaning?”. For “yes” responses, three follow-up scales probed distress, preoccupation, and conviction on a five-point scale (scored from 1 to 5). Delusion-like beliefs were scored as the sum of all responses to the dichotomous scale (scored as 0 for “no” and 1 for “yes”) and the three follow-up scales (internal consistency reliability for the dichotomous scale: \( \alpha = 0.88, \ \omega_f = 0.89 \) and \( \omega_h = 0.74 \), Revelle & Condon, 2019). For exploratory purposes, in addition to the PDI items, we included items from a scale that overlaps heavily with the PDI—the positive-symptoms
subscale of the Community Assessment of Psychic Experiences (CAPE; Stefanis et al., 2002). We do not analyze these items here.

*Cognitive Reflection Test (CRT; Frederick, 2005).* The CRT consists of verbal puzzles with intuitively appealing but incorrect answers (e.g., “A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?” The intuitive answer is $0.10; the correct answer is $0.05). We combined a three-item (Shenhav et al., 2012) and four-item CRT (Thomson & Oppenheimer, 2016). Analytic thinking was indexed as the number of correct answers (internal consistency reliability: $\alpha = 0.77$, $\omega_t = 0.80$, $\omega_h = 0.67$, Revelle & Condon, 2019). We also included four verbal reasoning items (Condon & Revelle, 2014) that are similar to the CRT in terms of content but do not have intuitively appealing incorrect answers and, thus, served as distractors. For exploratory purposes, after these questions were presented, we asked participants to rate on a five-point Likert scale whether they had seen any of these questions before.

*WORDSUM Test (Malhotra et al., 2007).* The WORDSUM test is a brief vocabulary test that consists of 10 target words in capital letters. Participants are asked to choose a word that most closely matches the meaning of the target word from five options. Data from this test were collected for exploratory purposes and are not analyzed here.

*Demographic questions.* Participants provided their age, gender, education level, and degree of belief in God. Finally, they indicated whether they were a native speaker of English and were asked if they had any comments about the study.

Each task was presented in the order listed above, except for the Beads Task and the BADE Task, which were always either first or second (order randomized). The order of trials and items in each task was randomized per participant. Table 1 summarizes the variables included in our pre-registered analyses.
### Table 1

**Description of the main variables**

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Task/scale name</th>
<th>Interpretation of higher score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic thinking</td>
<td>Cognitive Reflection Test (CRT)</td>
<td>More analytic cognitive style</td>
</tr>
<tr>
<td>Bias against disconfirmatory evidence (BADE)</td>
<td>BADE Task</td>
<td>More resistant to disconfirmatory evidence</td>
</tr>
<tr>
<td>Delusion-like beliefs</td>
<td>Peters et al. Delusions Inventory (PDI) total score</td>
<td>More delusion-like beliefs, and/or these beliefs endorsed with greater distress, conviction, and/or preoccupation.</td>
</tr>
<tr>
<td>Draws-to-decision</td>
<td>Beads Task</td>
<td>More beads drawn; less jumping to conclusions (JTC) bias</td>
</tr>
<tr>
<td>Liberal acceptance bias&lt;sub&gt;BADE&lt;/sub&gt;</td>
<td>BADE Task</td>
<td>More willing to endorse absurd explanations (<em>higher</em> liberal acceptance bias)</td>
</tr>
<tr>
<td>Liberal acceptance bias&lt;sub&gt;BEADS&lt;/sub&gt;</td>
<td>Beads Task</td>
<td>Higher subjective confidence when making decisions (<em>lower</em> liberal acceptance bias)</td>
</tr>
<tr>
<td>Overadjustment bias</td>
<td>Beads Task</td>
<td>Change in belief when a bead of the minority color is drawn, with a positive number indicating change towards the minority color (<em>higher</em> overadjustment)</td>
</tr>
</tbody>
</table>

### Statistical analyses

Unless otherwise noted, analyses reported in the main text were pre-registered. Our analyses depart from our pre-registration in one key respect: in place of frequentist regressions, we employ Bayesian regressions built with package *brms* (Bürkner, 2017) in R.
(R Core Team, 2023) to enable us to quantify the strength of evidence for null effects (Dienes, 2014; Wagenmakers et al., 2016). Unless otherwise indicated, ranges in parentheses represent 95% Credible Intervals (CrIs). Bayes Factors (BFs) represent the weight of evidence in favor of a given hypothesis (BF > 1) or in favor of a null effect (BF < 1), with BFs between 0.33 and 3 representing mere anecdotal evidence (although Bayesian approaches eschew dichotomous cut-offs).

**Data Quality**

As per our pre-registration, we examined whether our results were influenced by careless participants using two pre-registered screening methods. First, we included three attention checks and participants who failed two or more of them were classified as “careless”. In the BADE Task, the attention-check trial looked like the other BADE trials except that the final revealed statement read “To show you are being careful, please set all response sliders to the minimum value.” In the PDI task, the attention-check item read, “Have you ever used the internet? Answer ‘five’ for all three scales on the right.” In the CRT task, the attention-check item read “How many days are there in two weeks? To show you’re reading this, enter double the actual number.”

Second, participants who failed to move the sliders on more than a third of Beads Task and BADE Task questions that involved sliders were also classified as “careless”. In the Beads Task, questions that involved sliders were those that asked participants to rate how likely it was that one of two colors was the majority color of the urn. In the BADE Task, sliders were used to ask how probable the different scenarios were.

Participants who passed both screening methods were classified as “diligent”. We pre-registered that we would report results both with and without careless participants included in the sample, so as to examine the influence of careless participants on results. In addition to
the analyses reported below, in SM09 we carefully examine whether our attention checks really do identify “careless” participants rather than, say, participants who have unusual response patterns as a result of cognitive biases.

Results

Descriptive statistics and zero-order correlations are reported in SM02.

Data quality checks

8.2% of participants were classified as careless according to our pre-registered criteria because they either failed two or more attention checks, or did not move sliders on more than a third of tasks that included them. Figure 1 shows differences between careless participants and diligent participants in terms of the number of delusion-like beliefs endorsed (21 dichotomous PDI items), draws-to-decision in the Beads Task, plausibility rating of the two lure explanations in the BADE Task, and CRT score. These plots clearly demonstrate markedly different distributions of responses in careless participants compared to diligent participants. Moreover, as shown in SM02, in the graded-estimates version of the Beads Task, diligent participants typically decreased their confidence each time they saw a bead of the minority color; and across BADE Task stages the plausibility rating of the two lure responses decreased, the true response increased, and the absurd response was uniformly low. Thus, the overall patterns of responses among diligent participants were as expected.
Figure 1

Responses for Four Key Variables

Note. Responses are split by data quality (blue = diligent, red = careless). (a) Histogram of the total number of delusion-like beliefs endorsed (21 dichotomous PDI items). (b) Histogram of beads drawn in the draws-to-decision version of the Beads Task. (c) Mean ratings of plausibility for the two lure explanations for each stage of the BADE Task. (d) Number of items correct on the 7-item CRT. For summary statistics, see SM02.

Tests of hypotheses

The analyses under Research Questions 1 and 4 depart from our pre-registration (see footnotes 1 & 2). In the case of Research Question 4, the analysis below incorporates the pre-registered analysis, so we indicate when analyses expand on the pre-registration. For robustness checks supporting the results reported below, see SM03-07.

Research Question 1

Does analytic thinking predict the JTC bias and the BADE and, in turn, do these biases predict delusion-like beliefs?
**Hypothesis H1a.** Analytic thinking predicts the JTC bias and in turn, the JTC bias predicts delusion-like beliefs.

**Hypothesis H1b.** Analytic thinking predicts the BADE and in turn, the BADE predicts delusion-like beliefs.

For H1a, we regressed draws-to-decision (i.e., the inverse of JTC bias) on analytic thinking using a Bayesian zero-truncated negative binomial regression (for an account of this choice of model, motivated by the distribution of the draws-to-decision variable, see SM04). We simultaneously regressed delusion-like beliefs on draws-to-decision and analytic thinking using a zero-inflated negative binomial regression (again, see SM04). The resulting model is, in essence, a mediation model (Figure 2). With the full dataset, there was a positive association between analytic thinking and draws-to-decision ($b = 0.573 [0.457, 0.695]$, Figure 2 top left) and a negative association between draws-to-decision and delusion-like beliefs ($b = −0.023 [−0.033, −0.013]$). Participants higher in analytic thinking drew more beads, and participants who drew more beads were less likely to report delusion-like beliefs (supporting H1a). Even after accounting for draws-to-decision, analytic thinking had a direct negative association with delusion-like beliefs ($b = −0.682 [−0.875, −0.491]$) and a very small indirect negative association with delusion-like beliefs via draws-to-decision ($b = −0.013 [−0.019, −0.008]$).
Figure 2

Mediation models addressing Research Question 1

Note. Bayesian regression coefficients (with 95% Credible Intervals) regressing the cognitive biases draws-to-decision and the BADE (each in a separate model) on analytic thinking and simultaneously regressing delusion-like beliefs on analytic thinking and on the cognitive biases. The top row shows models with the full data set; the bottom row shows models with careless participants excluded. Pathways where the CrIs include 0 are in grey and labelled in italics.

For H1b, we used the same model structure as H1a, but replaced draws-to-decision with the BADE (Figure 2 top right). With the full dataset, there was a negative association between analytic thinking and the BADE \((b = -0.611 [-0.785, -0.293])\), and a positive association between the BADE and delusion-like beliefs \((b = 0.293 [0.238, 0.347])\).

Participants higher in analytic thinking were less biased against disconfirming evidence, and people who were more biased against disconfirming evidence were more likely to report delusion-like beliefs (supporting H1b). Even accounting for the BADE, analytic thinking had
a direct negative association with delusion-like beliefs ($b = -0.525 [-0.699, -0.347]$) and a smaller indirect negative association via the BADE ($b = -0.179 [-0.233, -0.129]$).

However, the results for both analyses changed when careless participants were excluded. Draws-to-decision was still positively—if more weakly—associated with analytic thinking ($b = 0.292 [0.180, 0.404]$, Figure 2 bottom left), but contra H1a draws-to-decision was no longer associated with delusion-like beliefs ($b = 0.004 [-0.007, 0.016], BF = 0.01$). BADE was still positively associated with delusion-like beliefs ($b = 0.081 [0.007, 0.155]$, Figure 2 bottom right), but contra H1b BADE was no longer associated with analytic thinking ($b = -0.022 [-0.176, 0.129], BF = 0.08$). In both cases, a direct association remained between analytic thinking and delusion-like beliefs (see Figure 2 bottom for details in each model), but in neither case was there an indirect association (via draws-to-decision: $b = 0.001 [-0.002, 0.005], BF = 0.005$; via the BADE: $b = -0.002 [-0.017, 0.012], BF = 0.012$). Thus, despite some non-null individual paths, there was no association from analytic thinking via these cognitive biases to delusion-like beliefs. For more detail on Research Question 1, including more complex path models, see SM03, SM04.

**Research Question 2**

Are the overadjustment bias and the BADE associated with each other and with delusion-like beliefs?

**Hypothesis H2a.** Delusion-like beliefs are not associated with at least one of these biases; *or*

**Hypothesis H2b.** Delusion-like beliefs are associated with both biases, but they are not themselves related; *or*

**Hypothesis H2c.** Delusion-like beliefs are associated with both biases, and they are themselves related.
With the full dataset, H2b was supported. There was a positive correlation between delusion-like beliefs and the overadjustment bias ($\rho_s = 0.108 \, [0.041, 0.172]$) and a positive correlation between delusion-like beliefs and the BADE ($\rho_s = 0.194 \, [0.129, 0.258]$, Figure 3), but there was strong evidence for no correlation between the BADE and the overadjustment bias ($\rho_s = 0.030 \, [-0.040, 0.101]$, $BF = 0.04$).

**Figure 3**

*Associations between Delusion-like Beliefs, Overadjustment and the BADE addressing Research Question 2*

Note. Spearman correlation coefficients ($\rho_s$) with 95% Credible Intervals.

This pattern appeared to change when careless participants were excluded (Figure 3). The correlations involving delusion-like beliefs were weaker (overadjustment bias $\rho_s = 0.060 \, [-0.008, 0.121], BF = 0.14$; BADE $\rho_s = 0.071 \, [0.008, 0.135], BF = 0.27$). In both cases there was moderate evidence ($0.1 < BF < 0.33$) for null relationships. The correlation between the BADE and the overadjustment bias remained null ($\rho_s = -0.042 \, [-0.107, 0.022], BF = 0.06$). For more detail on Research Question 2, see SM05.
Research Question 3

Do the JTC bias and the BADE each predict unique variance in delusion-like beliefs, and if so, which is the better predictor?

Hypothesis H3a. The JTC bias predicts delusion-like beliefs better than the BADE does; or

Hypothesis H3b. The BADE predicts delusion-like beliefs better than the JTC bias does; or

Hypothesis H3c. Neither bias is a better predictor of delusion-like beliefs.

We regressed delusion-like beliefs on draws-to-decision and the BADE using a Bayesian zero-inflated negative binomial model (for an account of this choice of model, see SM04). With the full dataset, there was a negative coefficient for draws-to-decision ($b = -0.089 [-0.137, -0.040]$; model $R^2 = 0.097 [0.062, 0.139]$) and a positive coefficient for the BADE ($b = 0.257 [0.209, 0.304]$, Figure 4). There was strong evidence ($BF > 1000$) that the coefficient for the BADE was larger in absolute size than that for draws-to-decision. Thus, in the full dataset, the BADE was more closely associated with delusion-like beliefs than the JTC bias was (supporting H3b).
Figure 4

Regression coefficients addressing Research Question 3

![Graph showing regression coefficients](image)

*Note.* Bayesian regression coefficients (with 95% Credible Intervals) for the predictor variables draws-to-decision and the BADE and the outcome variable delusion-like beliefs.

With careless participants excluded, draws-to-decision was centered on 0 ($b = 0.002 [-0.051, 0.055]$; model $R^2 = 0.004 [0.000, 0.012]$) and the coefficient for the BADE had a CrI that just included 0 ($b = 0.049 [-0.001, 0.100]$). There was moderate evidence ($BF = 8.52$) that the coefficient for the BADE was larger in absolute size than that for draws-to-decision. Thus, with careless participants excluded, the BADE was a marginally better predictor of delusion-like beliefs than the JTC bias was (again supporting H3b, although more weakly). For more detail on Research Question 3, see SM06.

**Research Question 4**

Are liberal acceptance bias\textsubscript{BEADS} and liberal acceptance bias\textsubscript{BADE} associated with each other and with delusion-like beliefs?
**Hypothesis H4a.** Liberal acceptance bias_{BEADS} and liberal acceptance bias_{BADE} are associated; and

**Hypothesis H4b.** Delusion-like beliefs are associated with at least one of these biases.

H4a was pre-registered; H4b was not, but it is important for our overall aim of understanding delusion-like beliefs. With the full dataset, liberal acceptance bias_{BEADS} and liberal acceptance bias_{BADE} were unrelated ($\rho_s = -0.046 [-0.109, 0.020], BF = 0.07$). The same was true with careless participants excluded ($\rho_s = -0.036 [-0.102, 0.025], BF = 0.05$). Thus, regardless of data quality, the liberal acceptance biases do not appear to capture the same (or a similar) underlying construct and might reflect task-specific strategies, *contra* H4a.

Regressing delusion-like beliefs on both versions of the liberal acceptance bias (using the same model family as for H3), analysis of the full data set indicated that liberal acceptance bias_{BEADS} was not predictive of delusion-like beliefs ($b = -0.045 [-0.096, 0.007], BF = 0.12$), but there was a positive coefficient for liberal acceptance bias_{BADE} with participants who rated the absurd explanations as more plausible also having more delusion-like beliefs ($b = 0.253 [0.209, 0.297], Figure 5$). Thus, analysis of the full dataset supports H4b. With careless participants excluded, neither variable was predictive of delusion-like beliefs (liberal acceptance bias_{BADE} $b = 0.035 [-0.017, 0.088], BF = 0.06$; liberal acceptance bias_{BEADS} $b = -0.025 [-0.076, 0.027], BF = 0.04$), *contra* H4b. For more detail on Research Question 4, see SM07.
**Discussion**

In the present study, we examined four broad research questions concerning relationships between cognitive biases (measured using the Beads Task and the BADE Task), delusion-like beliefs (measured using the PDI), and analytic thinking (measured using the CRT). Research Question 1 concerned relationships among analytic thinking, the most widely studied biases from the Beads Task and the BADE Task literature (the JTC bias and the BADE), and delusion-like beliefs. Irrespective of whether the full sample was used or careless participants were excluded, we found that analytic thinking was negatively associated with the JTC bias. While the JTC bias predicted delusion-like beliefs with the full sample, this relationship disappeared when careless participants were excluded. Conversely, analytic thinking was negatively associated with the BADE in the full sample, but there was...
no association when careless responders were excluded. However, the BADE predicted delusion-like beliefs regardless of whether or not careless participants were excluded. In all cases, analytic thinking was directly associated with delusion-like beliefs. Although there were indirect effects (via the JTC bias or the BADE) with the full sample, these were smaller than the direct effects, and both indirect effects disappeared once careless participants were excluded.

Research Questions 2, 3, and 4 concerned relationships among a suite of five cognitive biases and delusion-like beliefs. When including all participants in the analyses, we typically replicated classic findings from the literature: with the Beads Task, delusion-like beliefs were associated with the JTC bias and the overadjustment bias (but not the liberal acceptance bias); and with the BADE Task, delusion-like beliefs were associated with the BADE and the liberal acceptance bias. Strikingly, however, when careless participants were excluded, the picture changed dramatically: the extent to which cognitive biases predicted delusion-like beliefs was diminished severely; and in the case of the JTC bias, the overadjustment bias and the liberal acceptance bias, Bayes Factors provided moderate-to-strong evidence that there was no association.

Our study had four key strengths and innovations. First, we developed an animated Beads Task that facilitated comprehension of the task. Second, we included a series of data-quality checks to identify careless participants. Third, our sample is among the very largest in this literature and, to the best of our knowledge, is the single largest study that includes both the Beads Task and the BADE Task. Fourth, our hypotheses and analysis plan were pre-registered.

Strikingly, while only 8.2% of the sample were categorized as careless according to our pre-registered criteria, associations between delusion-like beliefs and cognitive biases
were severely attenuated—or disappeared entirely—when these careless participants were
excluded from analyses. Why? We argue the answer is to be found by considering how
careless participants respond to the PDI, the Beads Task, and the BADE Task.

First, the PDI elicits low mean scores for diligent participants because they typically
report holding few delusion-like beliefs (Peters et al., 2004). Consequently, if careless
participants respond randomly then, on average, they will be coded as having particularly
high levels of delusion-like beliefs compared to diligent participants. This is exactly what we
find: the mean number of items endorsed on the PDI (recall that participants answer “yes” or
“no” to each item) was 11 out of 21 for careless participants (close to the value of 10.5 that
would be expected if responses were entirely random) versus 3.7 for diligent participants; see
Figure 1a.

Second, in the Beads Task, if careless participants respond at random when deciding
whether to see an additional bead (or aim to finish the task quickly by asking to see few
beads) then, on average, they would ask to see fewer beads than diligent participants, which
would resemble a JTC bias. Again, this is exactly what we find: the mean number of beads
was 2.7 for careless participants versus 9.6 for diligent participants; see Figure 1b.

Third, in the BADE Task, if careless participants neglect to move sliders when
additional information about the scenario is revealed then their responses would be coded as
exhibiting a strong BADE. Indeed, among careless participants we found almost no change in
the plausibility rating for lure explanations across stages of the task, whereas diligent
participants decreased the plausibility rating as evidence accumulated; see Figure 1c.

Similar arguments can be made for the influence of random responding on the
overadjustment bias and the liberal acceptance biases – lack of sincere engagement with tasks
would be coded as exhibiting bias. Moreover, in the case of the CRT we see that careless
participants had a markedly lower mean number of correct responses compared to diligent participants (20% vs 58%); see Figure 1d.

What are the implications for the wider literature of our finding that key relationships all but vanish when careless participants are excluded? While data-quality checks are beginning to be used in clinical psychology research (Agley et al., 2022; Chandler et al., 2020; Moseley et al., 2021; Ophir et al., 2020; Zorowitz et al., 2021), they are still very rare in the literature which examines relationships between delusions, delusion-like beliefs and cognitive biases. Consequently, the problem of participant carelessness may run significantly deeper than the present study alone. In fact, we are aware of only two other studies with data-quality checks examining the relationship between delusion-like beliefs and cognitive biases (Ross et al., 2016; So et al., 2023) and no studies with data-quality checks in the case of clinical delusions.5 Ross et al. (2016) removed inattentive participants prior to analysis, and despite also having a large sample (n = 558) they, likewise, failed to find an association between delusion-like beliefs and the JTC bias. So et al. (2023) did not report statistics for overall PDI scores, but reported that participants who scored higher on the conviction subscale of the PDI showed reduced JTC, counter to predictions.

It is possible that the current study and the studies of Ross et al. (2016) and So et al. (2023) failed to find clear evidence of cognitive biases because they were conducted online, where participants can be careless (Brühlmann et al., 2020) and dishonest (Ross and Levy, 2023). However, this seems unlikely because research suggests that participants in online studies are at least as diligent—or even more diligent—than participants in the laboratory (Hauser & Schwarz, 2016; Kees et al., 2017; Ramsey et al., 2016). Moreover, as we

---

5 It should be noted that there are some examples from outside the delusions literature. A study by Bronstein, Everaert, et al. (2019) included data quality checks and found a relationship between paranoid ideation in the general population and performance in the BADE Task.
discussed earlier, there is strong evidence that participants in the laboratory frequently misunderstand the Beads Task (Balzan, Delfabbro, & Galletly, 2012; Balzan, Delfabbro, Galletly, et al., 2012; Howe et al., 2018). Thus, we propose that the existing Beads Task and BADE Task literature should be interpreted with considerable caution because participant inattentiveness or misunderstanding may have routinely generated spurious associations between delusions, delusion-like beliefs, and cognitive biases.6

Where do we go from here? In our opinion, a key direction for future research is to conduct new empirical studies that carefully examine participant attentiveness and task comprehension. The present study provides several examples of how this can be done. However, no single attention check or comprehension check could possibly capture all dimensions of inattention or miscomprehension (Brühlmann et al., 2020), and the approaches we have used in our study are no exception. Consequently, as has been argued in other literatures (Arthur et al., 2021), it seems clear to us that a diversity of attention checks and comprehension checks will be needed to triangulate on robust conclusions.

Until data quality is routinely assessed in studies with both clinical and non-clinical populations, it would be premature to draw strong conclusions about the theoretical status of continuum models of delusions rooted in cognitive biases (especially as there are many “moving parts”: how the cognitive biases are conceptualized; how cognitive biases are inferred from response patterns; how delusions and delusion-like beliefs are measured; how tasks are comprehended; and how all of these are impacted by careless responding).

Nonetheless, considering one plausible case as an example, we might speculate about what it would mean for psychosis continuum models if future research with robust data quality

---

6 We are not aware of any existing research that directly examines attentiveness and comprehension of the BADE Task.
checks were to find that the JTC bias and BADE are associated with delusions in clinical samples, but not with delusion-like beliefs in non-clinical samples. One interpretation could be that ostensibly “delusion-like” beliefs measured with instruments like the PDI are not, in fact, on a continuum with clinical delusions, as some scholars have argued (Lawrie et al., 2010; Parnas & Menriksen, 2016). Alternatively, perhaps delusion-like beliefs driven by anomalous experiences are a risk factor for developing clinical delusions that only manifest when coupled with reasoning biases (Coltheart & Davies, 2021; Furl et al., in press). Of course, we can only speculate about what future studies that carefully examine participant attentiveness and task comprehension might find, but the crucial point is that routine data quality assessment is necessary to make this literature robustly interpretable.

In conclusion, by showing that careless participants can be responsible for putative associations between delusion-like beliefs and cognitive biases, the current study contributes to an emerging literature which demonstrates the importance of verifying that participants respond diligently and that they understand tasks, to avoid spurious associations (Arthur et al., 2021), particularly when employing measures with skewed distributions of responses (Agley et al., 2022; Chandler et al., 2020; Moseley et al., 2021; Ophir et al., 2020; Zorowitz et al., 2021).

**Funding**

J.S. and R.M. were supported by the Cogito Foundation (grant number R10917), the British Academy (grant number SRG1819\190996) and the NOMIS Foundation (grant ‘Diversity in Social Environments’ for J.S., grant ‘Collective Delusions: Social Identity and Scientific Misbeliefs’ to R.M.). R.M.R. was supported by the Australian Research Council (grant number DP180102384), the John Templeton Foundation (grant number 62631), and a
Macquarie University Research Fellowship. The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References


Linscott, R. J., & van Os, J. (2013). An updated and conservative systematic review and meta-analysis of epidemiological evidence on psychotic experiences in children and adults: On the pathway from proneness to persistence to dimensional expression across mental disorders. *Psychological Medicine, 43*(6), 1133-1149. https://doi.org/10.1017/S0033291712001626


