Manuscript Title:

Unstable belief-formation and slowed decision-making: evidence that the JTC reasoning bias in schizophrenia is not linked to impulsive decision-making.

Running Title: JTC in SCZ not linked to impulsive decision-making.

Authors and affiliations:

Wolfgang Strube[1,2], Camelia Cimpianu[1], Miriam Ulbrich[1], Ömer Öztürk[1,3], Thomas Schneider-Axmann [1], Peter Falkai[1], Louise Marshall[4], Sven Bestmann[4,5*], Alkomiet Hasan [2*]

[1] Department of Psychiatry and Psychotherapy, Ludwig Maximillian University, Munich, Germany
[2] Department of Psychiatry, Psychotherapy and Psychosomatics of the University Augsburg, Medical Faculty, University of Augsburg, Bezirkskrankenhaus Augsburg, Augsburg, Germany
[3] International Max Planck Research School for Translational Psychiatry, Munich, Germany
[4] Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, Queen Square, London
[5] Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, Queen Square, London

[(*)] both authors contributed equally to this work

Corresponding Author:
Wolfgang Strube, MD
Department of Psychiatry, Psychotherapy and Psychosomatics,
Medical Faculty University of Augsburg, BKH Augsburg
Dr. Mack-Straße 1, D-86156 Augsburg, Germany
Phone: +49 821 4803 1011
Fax: +49 821 4803 1012
Email: wolfgang.strube@bkh-augsburg.de

Word and reference count

Abstract	249	(maximum 250)
Article	3765	(maximum 4000)
No. of References	57	(no limit)
Tables & Figures	4	(maximum 5)
Supplements	2	(no limit)

Abstract

Background: Jumping-to-conclusions (JTC) is a prominent reasoning bias in schizophrenia (SCZ). While it has been linked to psychopathological abnormalities (delusions and impulsive decision-making) but also unstable belief-formation, its origin remains unclear. We here directly test to which extend JTC is associated with delusional ideation, impulsive decision-making, and unstable belief-formation.

Methods: 45 SCZ patients were compared to matched samples of 45 patients with major depressive disorder (MDD), and 45 healthy controls (HC) as delusions and JTC also occur in other mental disorders and the general population. Participants performed a probabilistic beads task. To test the association of JTC with measures of delusions ($PANSS_{positive}$, $PANSS_{positive\text{-}factor}$, PDI), Bayesian linear regressions were computed. For the link between JTC and impulsive decision-making and unstable beliefs we conducted between-group comparisons of 'draws-to-decision' (DTD), 'decision times' (DT), and 'disconfirmatory evidence scores' (DES).

Results: Bayesian regression obtained no robust relationship between PDI and DTD (all $|R^2_{adj}| \leq 0.057$, all $p \geq 0.022$, all $BF_{01} \leq 0.046$; $\alpha_{adj} = 0.00833$). Compared to MDD and HC, patients with SCZ needed more time to decide (significantly higher DT in ambiguous trials: all $p \leq 0.005$, $r^2 \geq 0.216$; numerically higher DT in other trials). Further, SCZ had unstable beliefs about the correct source jar whenever unexpected changes in bead-sequences (disconfirmatory evidence) occurred (compared to MDD: all $p \leq 0.004$, all $r^2 \geq 0.232$; compared to HC: numerically higher DES). No significant correlation was observed between decision-times and draws-to-decision (all $p \geq 0.050$).

Conclusions: Our findings point towards a relationship of JTC with unstable belief formation and do not support the assumption that JTC is associated with impulsive decision-making.

Keywords: Schizophrenia – beads task – jumping-to-conclusions (JTC) bias – probabilistic reasoning – unstable belief-formation – slowed decision-making

Introduction

People with schizophrenia (SCZ) often gather less information before arriving at a conclusion [1-6]. This jumping-to-conclusions (JTC) bias is most commonly assessed with the beads task [7], which requires participants to sample random sequences of colored beads and to infer from which of two potential source jars the beads were taken. Essentially, the prediction of the beads task is that patients with SCZ will view fewer beads before deciding, and commonly do so with greater confidence than controls. Two variants of the beads task have commonly been used to assess JTC in SCZ. In the so called draws-to-decision version [5, 7, 8], participants are allowed to sample any desired number of up to 20 beads and following every new bead either continue sampling, or decide from which of the two source jars the beads were drawn – as soon as they feel sufficiently certain to make this decision. In the graded-estimates version [9, 10], participants are shown a fixed number of beads and asked on each new bead view to rate the probability (on a Likert scale ranging from 0% to 100%) that the currently viewed bead sequence is drawn from one of the two source jars.

Patients with SCZ consistently show JTC [1, 4, 11] and over-estimate the probability that the beads are drawn from a specific jar [9, 10, 12-14]. However, the reasons for this bias remain unclear [1, 4, 11, 15-18]. On the one hand, JTC has been associated with delusions [1-4, 11, 19-23] and impulsive decision-making [1, 2, 22, 24]. On the other hand, a growing number of studies have shown that patients with SCZ display JTC even without the presence of delusions and other authors have attributed JTC to impaired probabilistic reasoning [10, 11, 16, 18, 25-27]. Specifically, patients seem to overweight unexpected recent bead occurrences – leading to unstable beliefs about the correct source jar [9, 12, 15, 17, 28]. Resolving these disparate views is of relevance as the JTC bias has become important for evaluating behavioral treatment interventions aimed at improving impulsive decision-making and adhering to inordinately fixed beliefs or aberrant reasoning [29, 30].

Against this background, we here directly assess the relationship of JTC with delusions, impulsive decision-making, and belief-formation. For this purpose, we applied a novel probabilistic beads task that allows

systematic assessments of data-gathering and explicit probability estimates from varying sensory and cognitive information (via different trial-types) and in different probabilistic contexts (via different source jar distributions). We hypothesized (a) that JTC is more pronounced in SCZ patients compared to matched clinical and healthy controls. Further, we reasoned that JTC is (b) not strongly associated with delusions or (c) impulsive decision-making. Finally, we hypothesized (d) that JTC would be associated with specific alterations of probabilistic reasoning, namely unstable belief-formation.

Methods

Study population

Inpatients currently treated at the department of psychiatry of the University Hospital Munich (Germany) diagnosed with schizophrenia (SCZ, n=45) or major depression disorder (MDD, n=45) participated in this study. Diagnoses were based on non-structured clinical interviews following ICD-10 criteria and were confirmed by two independent clinical interviewers. Additionally, n=45 healthy control participants (HC) were tested. Both groups were matched to SCZ participants (see Supplements 'Methods') with respect to age, gender, and intelligence quotient (IQ) as these factors have been identified to impact beads task performance [20, 31]. Further, clinical samples were matched according to disease duration. We included a clinical control group of MDD patients, because dysfunctions in decision-making are considered a central psychopathological feature of depression and to remain comparable with previous studies on MDD patients [1, 4, 31, 32]. Our sample size met the requirements of a corresponding power analysis (G*Power, [33]) assuming a small expected effect size of f = 0.25 an alpha error probability of $\alpha = 0.05$, a power of (1-β err prob) = 0.9, and a correlation among repeated measures of 0.5 (see Supplements 'Methods'). Inclusion and exclusion criteria and procedures for obtaining clinical and neuropsychological characteristics are detailed in Supplement 'Methods'. All participants provided informed consent in accordance with the standards of the Declaration of Helsinki and the study protocol was approved by the local Ethics Committee. Monetary compensation amounted to 20€ per testing hour. In case of SCZ and HC, available

data sets [34] were extended with newly recruited participants from n=32 to n=45 each. Symptom severity and global functioning were assessed via Clinical-Global-Impression (CGI) [35] and Global-Assessment-of-Functioning (GAF) [36]. For SCZ, we further conducted the Positive-and-Negative-Syndrome-Scale (PANSS) [37] and computed Wallworks' $PANSS_{positive\text{-}factor}$ (consisting of items P1, P3, P5, and G9) [38]. In case of MDD, we surveyed Beck-Depression-Inventory (BDI) [39] and Hamilton-Major-Depression-Rating-Scale (H-MDRS) [40]. All participants further underwent Peter-Delusions-Inventory (PDI) [41] and neuropsychological assessments of attention [42], executive functioning [43]), and premorbid intelligence [44].

Behavioral Tasks

Upon inclusion, all participants performed the same set of behavioral reach decision tasks that were computerized using Matlab® (Mathworks, Massachusetts, version R2017b) and the Cogent Graphics toolbox for Windows [45]: first, the Choice Reaction-time Task (CRT), followed by our computational instantiation of the beads task [7, 10, 46] (see Figure 1).

Choice Reaction-time Task (CRT)

The CRT was used to obtain individual reaction times (RT_{CRT}), which are considered [47] aggregates of the delays attributable to individual differences in attention span and sensory processing, muscle response initiation, and cognitive slowing. As in previous work [47], we corrected the response times in the beads task ($RT_{Beadstask}$) for these aggregates of delays on an individual subject basis (correction = $RT_{Beadstask}$ - RT_{CRT}) to estimated individual 'decision times' (DT), as a proxy for the amount of time participants spent to cognitively process bead sequences and make a decision.

Fig. 1

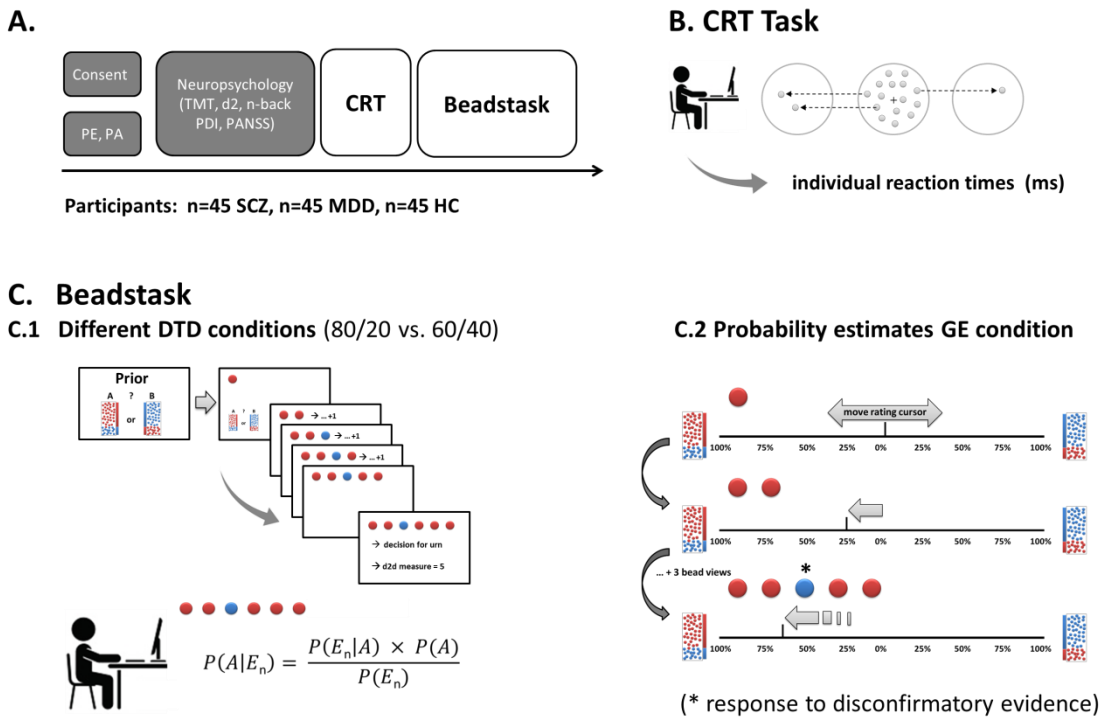

Figure 1 A: Overview of the experimental procedure. In total n=45 patients with schizophrenia (SCZ), n=45 patients with major depressive disorder (MDD), and n=45 healthy controls (HC) participated. **B**: Visualizations of the Choice Reaction-time Task (CRT), which obtained individual reaction times. **C**: Overview of the beads task design: **C.1** Procedure for the two draws-to-decision (DTD) conditions. Two sets of different colors were used to facilitate distinguishing between the two conditions. **C.2** The graded estimates (GE) condition was used to obtain probability-estimates (PE) and disconfirmatory evidence scores (DES). DES quantified the cumulative amount by which participants and patients changed their PE following changes in bead color after viewing ≥2 preceding beads of the same color (i.e. following disconfirmatory evidence). DES scores thus quantified participants' responses to surprising (i.e. disconfirmatory) evidence caused by a presenting beads of the color opposite to the participants' belief about the predominant bead color in the presumed source jar. **C.3** Bead sequences were generated with three different likelihoods [$P(A|E_n)$] for the three respective trial-types: (I) easy trials with 80% likelihood for one predominant bead color, (II) difficult trials with 60% likelihood, and (III) ambiguous trials with 50% likelihood.

Note: *P(A)* denotes the probability for jar 'A' being correct in each new trial; $P(E_n|A)$ represents the likelihood of each bead sequence to be drawn from jar 'A' (E: sensory evidence [current color sequence of beads presented], n: number of beads); $P(E_n)$ denotes the total probability of the given data and $P(A|E_n)$ the inferred likelihood of jar 'A' to be correct given the sensory evidence E_n accumulated. See [34] for methodological details on the graded-estimates (GE) version, which was used to obtain explicit probability estimates (PE) that participants rated on a Likert scale for presented bead sequences.

Beads Task

The complete task design is detailed in the supplemental 'Methods'. For the beads task, we applied two draws-to-decision (DTD) conditions (2×18 trials), which differed regarding the prior information participants were given about the distributions in each jar ($P_{80/20}$: 80:20%, blue:green; $P_{60/40}$: 60:40%, violet:orange). Here, participants could stop viewing additional beads whenever they felt sufficiently confident to make a decision about a source jar. Additionally, participants undertook 12 trials of the graded-estimates (GE) condition, which also used a $P_{80/20}$ distribution of beads. In the GE, participants could not terminate a trial after viewing a bead, but instead were required to view ten beads successively and to report probability estimates (PE) on a Likert-scale (ranging from 0% to 100% probability in two directions, with both source jars presented at the extremes). These PE reflected how likely participants estimated a bead sequence originated from a source jar they selected (participants could decide for either of the two potential source jars, but not for both). As in previous work [12, 34], disconfirmatory evidence scores (DES) were further assessed that quantified the cumulative amount by which participants changed their PE following changes in bead color after viewing ≥2 preceding beads of the same color. Participants were instructed to view and rate random bead sequences, however, pre-specified trial sequences were applied [34, 46]: (I) easy trials with a likelihood of 80% for one predominant bead color, (II) difficult trials with a likelihood of 60%, and (III) ambiguous trials with a likelihood of 50%. All sequences were counterbalanced and the order of trials was randomized to control for sequence effects. In all three parts of the task, no feedback was provided about the correctness of responses. Bead sequences and illustrations of the source

jars were kept displayed on the computer screen throughout the whole task to reduce working memory load, which has been demonstrated to bias beads task findings [16, 19, 48].

Statistical Analyses

All statistical analyses were conducted using IBM SPSS 27. Level of significance was set to alpha=0.05 for group level comparisons of baseline characteristics using one-way analysis of variance (ANOVA) and Chi-square tests. We defined DTD, JTC frequencies, DT, PE, and DES as main outcome variables. As the assumption of normal distribution was violated for 52 of the 53 metric main outcome variables (Shapiro-Wilk-tests: 5th PE rating in easy trials: $W_{(135)}$=0.982, p=0.074; all other $W_{(135)}$≤0.980, all p≤0.048) and several methods of transformation did not achieve normal distribution, between-group comparisons (SCZ/MDD/HC) were computed using non-parametric Kruskal-Wallis tests (KWT) and - where appropriate - corrected Mann-Whitney-U tests (MWU), instead of repeated measures ANOVA assumed for our power analysis (see Supplement 'Methods'). For hypothesis (a) JTC frequencies were compared with Freeman-Halton tests [49] (in analogy of Chi-Square tests as contingency tables measured 2x3 and some expected cell counts were < 5). Data-gathering was investigated comparing DTD between groups using KWT and MWU. Hypothesis (b) was assessed with Bayesian linear regression between DTD and measures of delusions (PDI, $PANSS_{positive}$, $PANSS_{positive\text{-}factor}$ [38]) and age, gender, IQ and disease severity included as co-factors. For hypothesis (c) we compared DT between groups using KWT and MWU. To assess hypothesis (d) KWT and MWU were computed comparing PE and DES between groups. Effect sizes were estimated using $\Phi_c=(\sqrt{\chi^2/\sqrt{n})}$ for Freeman-Halton-tests, $\eta^2=(H-k+1)/(n-k)$ for KWT, $r^2=(Z^2/n)$ for MWU, R^2 for Bayesian regression [50]. To correct for the number of tests per main outcome analysis, we adjusted the significance level to α_{JTC}=0.00833, α_{DTD}=0.00333, $\alpha_{Bayesian\text{-}Regression\text{-}Delusions}$=0.00833, α_{DT}=0.00833, α_{PE}=0.00125, and α_{DES}=0.01667 (see Supplement 'Methods', section 1).

Results

Baseline characteristics across groups

Participants were well matched regarding group distributions of age, gender, IQ, and attention load capacity (see Supplement 'Results', section 1 and Table 1). SCZ and MDD displayed gradually slower response times in the CRT test compared to HC. Clinical characteristics further categorized symptom burden in both patient groups as moderate to severe [51, 52] and psycho-social level of function as higher in MDD compared to SCZ (see Table 1). Post-hoc analyses of individual responses on a trial-by-trial basis obtained no indicators of non-comprehension or reduced motivation (e.g. uniform responses). All participants were able to adequately differentiate between trial-types and conditions (see Supplement 'Results', section 1 and Tables 1-6).

JTC frequencies and data-gathering

To assess our hypothesis (a) whether SCZ patients showed increased rates of JTC we first classified JTC as present according to established criteria [1, 3, 4, 53], i.e. if participants decided after < 3 bead-views in the draws-to-decision conditions ($P_{80/20}$/$P_{60/40}$). For the $P_{80/20}$ condition, we observed JTC frequencies of 44.4% (n=20) in easy trials, 40.0% (n=18) in difficult trials and 44.4% (n=20) in ambiguous trials for SCZ patients. By comparison, JTC was less frequently displayed by MDD (easy trials: 8.9%, n=4; difficult trials: 6.7%, n=3; ambiguous trials: 4.4%, n=2) and by HC (easy trials: 13.3%, n=6; difficult trials: 13.3%, n=6; ambiguous trials: 2.2%, n=1). Freeman-Halton tests comparing these distributions for each trial-type confirmed the significance of these observed differences (overall analysis: all $X^2_{(2)} \geq 16.4$, all $p \leq 0.001$, $\Phi_c \geq 0.360$) and obtained higher JTC rates in SCZ compared to MDD (all $X^2_{(1)} \geq 13.98$, all $p < 0.001$, all $\Phi_c \geq 0.394$) and compared to HC (all $X^2_{(1)} \geq 8.18$, all $p \leq 0.008$, all $\Phi_c \geq 0.302$), while MDD and HC showed no differences (all $X^2_{(1)} \leq 1.11$, all $p \geq 0.242$). In contrast, SCZ displayed JTC less frequently in the $P_{60/40}$ condition (easy trials: 15.6%, n=7; difficult trials: 11.1%, n=5; ambiguous trials: 6.7%, n=3) and we observed no significant group differences (all $p \geq 0.507$, see Supplement 'Results', Tables 7-8).

Tab. 1

Group	SCZ	MDD	HC	statistics	
Demographics	**all (n=45)**	**all (n=45)**	**all (n=45)**	X^2 **(df)**	***p***
Gender (female : male)	19 : 26[b]	26 : 19[b]	22 : 23	2.19 (2)	0.334
	m (sd)	**m (sd)**	**m (sd)**	**F (df_1, df_2)**	***p***
Age	37.3 (11.9)	37.6 (11.3)	37.9 (11.4)	0.03 (2,132)	0.972
Severity of illness	**m (sd)**	**m (sd)**	**m (sd)**	**F (df_1, df_2)**	***p***
Disease duration (years)	10.1 (6.9)	11.3 (6.3)	-	[#]0.67 (1,86)	0.414
CPZ	328.7 (316.7)	-	-		
$PANSS_{positive}$	21.0 (5.8)	-	-		
$PANSS_{negative}$	17.1 (5.8)	-	-		
$PANSS_{general}$	38.3 (9.7)	-	-		
$PANSS_{total}$	76.4 (18.0)	-	-		
$PANSS_{positive\text{-}factor}$ [a]	13.1 (3.8)				
H-MDRS	-	20.7 (6.8)	-		
BDI	-	24.5 (10.0)	-		
GAF	58.5 (10.0)	63.6 (5.2)	-	[#]**9.48 (1,87)**	**0.003**
CGI	4.3 (0.5)	4.3 (0.7)	-	[#]0.35 (1,87)	0.557
PDI	31.8 (9.1)	6.2 (7.4)	1.6 (2.4)	**248.2 (2,132)**	**<0.001**
Neuropsychological Tests	**m (sd)**	**m (sd)**	**m (sd)**	**F (df_1, df_2)**	***p***
Premorbid IQ (PIA-IQ)	106.6 (3.3)	107.3 (4.3)	106.9 (3.5)	[#]0.43 (2,130)	0.654
CRT reaction times (msec)					
First run	429.4 (151.9)	409.4 (138.8)	384.6 (87.6)	1.36 (2,132)	0.260
Second run	430.7 (154.0)	416.3 (149.2)	380.4 (79.8)	1.73 (2,132)	0.181
TMT A performance score (sec)	40.7 (20.6)	26.1 (7.0)	26.4 (11.3)	**15.60 (2,132)**	**<0.001**
TMT B performance score (sec)	91.5 (35.0)	59.2 (20.1)	65.3 (21.3)	**19.09 (2,132)**	**<0.001**
d2 attention task score	233.0 (67.2)	248.2 (44.8)	252.5 (36.4)	[#]1.75 (2,129)	0.177

Table 1 Sociodemographic, clinical and neuropsychological characteristics. Statistics reflect group comparisons. Index: n: number of participants; m: mean; sd: standard deviation; df: degrees of freedom; X^2: Chi-square test; F: F-statistic; msec: milliseconds; PANSS: positive and negative syndrome scale; $PANSS_{positive}$: PANSS positive subscale score; $PANSS_{negative}$: PANSS negative subscale score; $PANSS_{general}$: PANSS general subscale score; $PANSS_{total}$: PANSS total sum score; [a] $PANSS_{positive\text{-}factor}$ according to Wallwork et al. 2012; H-MDRS: Hamilton disease rating scale for depression; BDI: Beck depression inventory; GAF: global assessment of functioning; CGI: clinical global impression scale; PDI: Peters et al. delusions inventory; IQ: intelligence quotient, PI A-IQ: premorbid intelligence assessment of IQ; CRT: Choice Reaction-time Task; TMT: trail making task; [#] indicate missing values in 1 (GAF, CGI), 2 (disease duration, premorbid IQ), or 3 (d2 attention task) participants, respectively. Significant results are highlighted in bold. [b] Note: the observed divergent gender distributions are common for studies on patients with schizophrenia (usually more male participants) and patients with major depression (usually more female participants).

We next compared draws-to-decisions (DTD) in each trial-type between groups since the definition of JTC (at < 3 bead views) does not consider the composition of the first two bead views in a sequence. This can largely impact the probability estimates, as the first two beads can be of the same color (and point into the direction of one specific source-jar) or not. For the $P_{80/20}$ condition respective KWT obtained significant group differences in easy ($H_{(2)}$=18.43, p<0.001, η^2_H=0.391), difficult ($H_{(2)}$=15.87, p<0.001, η^2_H=0.330), and ambiguous trials ($H_{(2)}$=19.04, p<0.001, η^2_H=0.406). Subsequent Sidak corrected MWU explained the observed effects through significantly fewer DTD of SCZ patients compared to MDD patients (all U≤541.5, all p<0.001, all r^2≥0.321) and numerical differences compared to HC participants (all U≤685.0, all p≤0.024, all r^2≥0.155, α_{adj}=0.0033), while no group differences were observed between MDD and HC (all U≥796.5, all p≥0.240). By contrast, no significant differences were observed for the $P_{60/40}$ condition (all $H_{(2)}$≤5.14, all p≥0.077) (see Figure 2A and Supplement 'Results', Tables 5-6).

Delusions and JTC

Regarding our second hypothesis (b) whether JTC is linked to delusions, we computed Bayesian linear regression between PDI scores and draws-to-decision and included age, gender, IQ and disease severity as co-factors into our model. Since the PDI is considered transdiagnostic to assess delusional thinking across a continuum [41, 54], we first conducted a cross-sectional analysis on all participants (SCZ, MDD, and HC combined). This analysis obtained no significant correlations for all trial-types of the $P_{80/20}$ condition (all $|R^2_{adj}|$≤0.057, all $F_{(4,128)}$≤2.98, all p≥0.022, all BF_{01}≤0.046; α_{adj}=0.00833) and for the $P_{60/40}$ condition (all $|R^2_{adj}|$≤0.007, all $F_{(4,128)}$≤1.23, all p≥0.302, all BF_{01}≤0.002). Next, the same analysis was repeated separately for SCZ patients (only with $PANSS_{positive}$ instead of PDI) which showed no significant correlations (all $|R^2_{adj}|$≤0.057, all $F_{(4,40)}$≤0.609, all p≥0.658, all BF_{01}≤0.004). Similarly, no significant correlations were observed between Wallworks' $PANSS_{positive\text{-}factor}$ [38] and DTD (all $|R^2_{adj}|$≤0.057, all $F_{(4,40)}$≤0.602, all p≥0.663,

all BF_{01}≤0.004). However, our sample size was only sufficiently powered to detect PDI differences across groups, as secondary analyses for PDI differences within groups using G*Power [33] obtained larger necessary sample sizes (see Supplement 'Results', section 2).

Fig. 2

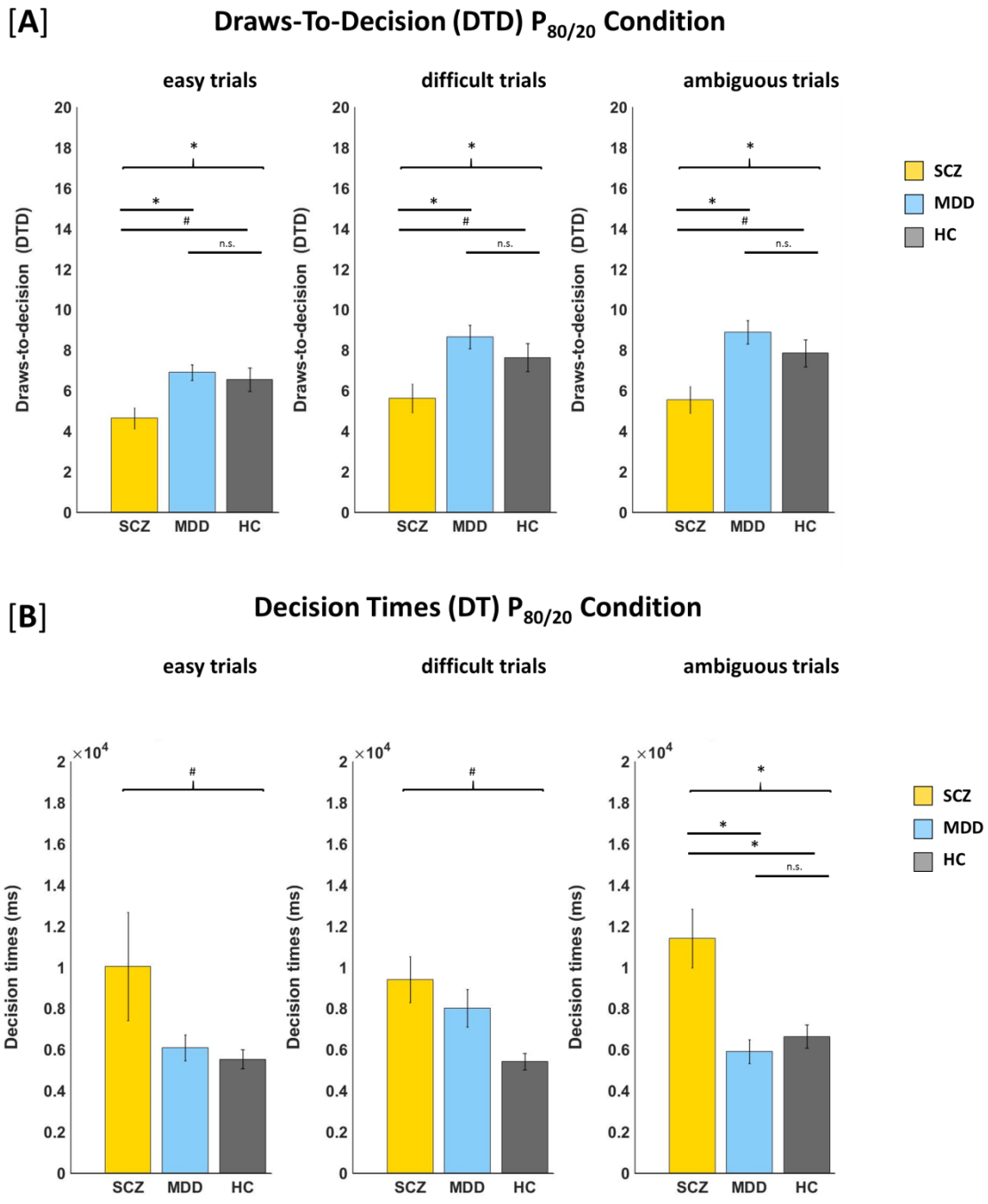

Figure 2

A Group-wise comparisons of draws-to-decision (DTD) in the $P_{80/20}$ condition subdivided by trial difficulty (easy – difficult – ambiguous trial-types). SCZ patients sampled less beads compared to MDD or HC in the context of the $P_{80/20}$ jar distribution. Error bars represent standard error of the mean. n.s.: not significant, *: $p<0.05$ in Kruskall-Wallis tests and post-hoc Sidak corrected Mann-Whitney-U tests, #: trend-level differences. Adjusted p-level: $\alpha_{adj}=0.00333$.

Figure 2

B Group-wise comparisons of decision times (DT) [measured in milliseconds (ms)]. As suggested in previous work [47], we corrected the response times in the beads task for aggregates of delays in attention span and sensory processing, muscle response initiation, and cognitive slowing to estimate the here reported individual 'decision times' (DT), as a proxy for the amount of time participants spent to cognitively process bead sequences and make a decision (see Methods section for further details). In both task conditions ($P_{80/20}$ and $P_{60/40}$) patients with SCZ needed more time to make decisions compared to MDD and HC (significantly higher DT in ambiguous trials: all $p\leq0.005$, $r^2\geq0.216$; numerically higher DT in other trials). Error bars represent standard error of the mean. n.s.: not significant, *: $p<0.05$ in Kruskall-Wallis tests and post-hoc Sidak corrected Mann-Whitney-U tests, #: trend-level differences. Adjusted p-level: $\alpha_{adj}=0.00833$.

Decision times and JTC

To investigate our hypothesis (c) whether JTC is associated with impulsive decision-making we compared individual decision times (DT) derived from reaction times of the CRT offset against beads task reaction times (as detailed in Methods). Respective KWT obtained significant group differences for ambiguous trials of both conditions ($P_{80/20}$ condition: $H_{(2)}=19.7$, $p<0.001$, $\eta^2_H=0.421$; $P_{60/40}$ condition: $H_{(2)}=11.5$, $p=0.003$, $\eta^2_H=0.227$), while only numerically slower DT were observed for easy and difficult ~~for~~ trials in both conditions (all $H_{(2)}\geq6.87$, all $p\leq0.032$, all $\eta^2_H\leq0.148$, $\alpha_{adj}=0.00833$). To further explore the direction of these differences, we further computed (in part exploratory) Sidak corrected MWU for the $P_{80/20}$ condition. This analysis obtained no differences between MDD and HC (all $U\geq813.0$, all $p\geq0.317$), while slower DT were observed in ambiguous trials comparing SCZ patients to MDD ($U=488.0$, $p<0.001$, $r^2=0.398$) and to HC

(U=626.0, p=0.005, r^2=0.216) (see Figure 2B). For the $P_{60/40}$ condition we observed significantly slower DT in ambiguous trials of SCZ compared to MDD (U=606.0, p=0.003, r^2=0.239). Interestingly, SCZ patients also showed a pattern of numerically slower decision times in easy and difficult trials of both conditions ($P_{80/20}$ and $P_{60/40}$) compared to MDD and HC (Supplement 'Results', section 3, Tables 9-10).

Probability ratings and unstable beliefs

In case of the graded estimates (GE) version of the beads task group comparisons using KWT and Sidak corrected MWU obtained a pattern of significantly and numerically higher probability estimates (PE) of SCZ patients in easy and difficult trials compared to MDD and HC, while no such differences were observed between MDD and HC (see Figure 3A and Supplement 'Results', section 4 and Tables 11-13). Further exploratory analyses obtained significant Spearman correlations between DES and DTD in easy (all $|r_S|\geq0.303$, all $p<0.001$) and difficult trials (all $|r_S|\geq0.242$, all $p\leq0.005$; $\alpha_{adj}=0.006$), but not for ambiguous trials (all $p\geq0.014$).

Further, with respect to our hypothesis (d) whether JTC is associated with unstable belief-formation, we computed 'disconfirmatory-evidence-scores' (DES), which quantify changes of probability ratings following switches of bead color compared to ≥2 preceding beads [12]. This approach was based on previous studies observing that patients with SCZ display unstable beliefs following unexpected changes in a bead sequences [12, 34, 55-59]. Respective KWT between SCZ, MDD, and HC obtained significant group differences for easy ($H_{(2)}$=17.87, $p<0.001$, η^2_H=0.237), difficult ($H_{(2)}$=22.19, $p<0.001$, η^2_H=0.233), and ambiguous trials ($H_{(2)}$=11.81, p=0.003, η^2_H=0.142). Subsequent Sidak corrected MWU explained the observed effects through significantly higher DES of SCZ patients compared to MDD patients in easy and difficult trials (all $U\leq612.0$, all $p\leq0.004$, all $r^2\geq0.232$; ambiguous trials: U=746.0, p=0.092), while only numerically higher DES were observed in difficult and ambiguous trails of SCZ patients compared to HC (all $U\leq678.5$, all $p\geq0.020$; easy trials: U=748.0, p=0.097). Further, no group differences were observed between

MDD and HC (all U≥874.0, all p≥0.778) (see Figure 3B). Finally, an exploratory analysis did not obtain a significant correlation between DES and decision times (all $|r_S|$≤0.169, all p≥0.050, all BF_{01}≥1.736).

Fig. 3

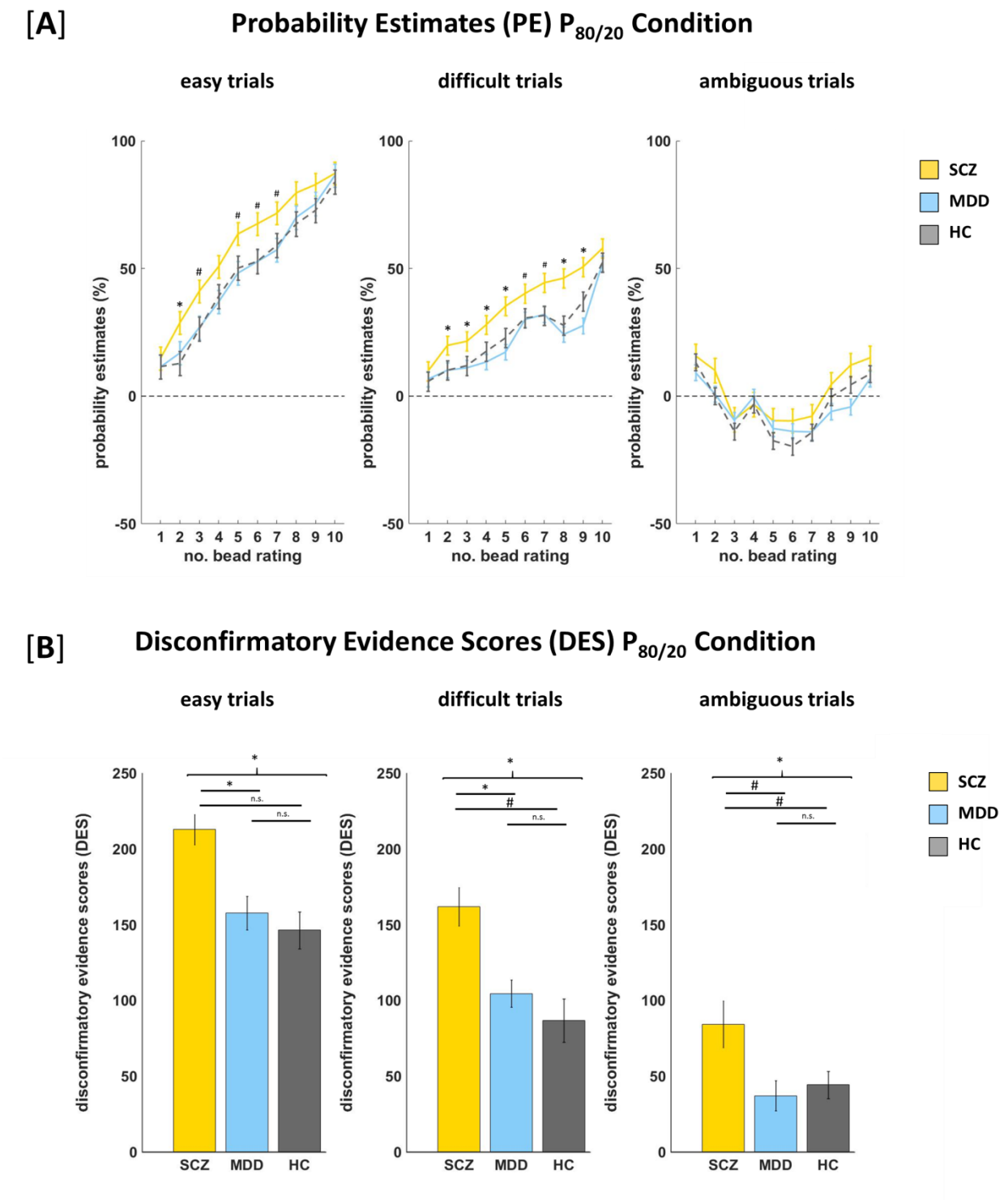

Figure 3

A Group-wise representation of mean probability estimates (PE) in the graded-estimates (GE) version of the task subdivided by trial difficulty (easy – difficult – ambiguous trial-types). SCZ patients showed a pattern of significantly and numerically higher probability estimates (PE) in easy and difficult trials compared to MDD and HC, while no such differences were observed between MDD and HC. Error bars represent standard error of the mean. Reference lines at 0% likelihood indicate 0% probability for either of the two source jars. *: $p<0.05$ in Kruskall-Wallis tests for each of the 10 ratings, #: trend-level differences. Adjusted p-level: $\alpha_{adj}=0.00125$.

B Group-wise comparisons of disconfirmatory evidence scores (DES) subdivided by trial difficulty (easy – difficult – ambiguous trial-types). DES quantified the cumulative amount by which participants and patients changed their PE following changes in bead color after viewing ≥2 preceding beads of the same color (i.e. following disconfirmatory evidence). DES scores thus quantified participants' responses to surprising (i.e. disconfirmatory) evidence caused by presenting beads of the color opposite to the participants' belief about the predominant bead color in the presumed source jar. SCZ patients displayed increased DES in each trial-type. Error bars represent standard error of the mean. n.s.: not significant, *: $p<0.05$ in Kruskall-Wallis tests and post-hoc Sidak corrected Mann-Whitney-U tests, #: trend-level differences. Adjusted p-level: $\alpha_{adj}=0.01667$.

Discussion

As main findings we here observed significantly higher rates of JTC and reduced DTD in patients with SCZ compared to MDD and HC and that SCZ patients displayed unstable belief-formation following unexpected changes in bead sequences. Our results also show for the first time that SCZ patients needed more time to view less beads before a decision.

The lower draws-to-decision (DTD) and correspondingly increased rates of JTC in SCZ patients compared to MDD and HC (see Figure 2A) side with recent meta-analyses [1-4]. By contrast, we did not find support for the hypothesis that these findings are also associated with delusions. This is in line with previous reports [26, 27] that challenge the view that JTC may reflect delusion-specific alterations. Of note in this regard, subsequent analyses showed that our sample size was too small to detect small effect sizes (as reported

in [60]) for associations of DTD with PDI; additionally, our sample of SCZ patients showed similar levels of delusional ideation, which probably precluded us from obtaining correlations of $PANSS_{positive}$ or $PANSS_{positive\text{-}factor}$ with DTD in our Bayesian linear regression. Due to these limitations the association between JTC and delusions remains unclear.

However, with respect to the proposed association of JTC with impulsive decision-making, group wise comparisons of decision times (DT; calculated specifically for this purpose) showed differences in ambiguous trials and numerical differences in easy and difficult trials but in an opposite way as previously assumed: although SCZ patients tended to decide after fewer bead views they showed a pattern of needing more time than MDD and HC to make their decisions. While we are aware that our approach to correct for contributors to slower reaction times in patients with SCZ (see Methods section and [47]) poses only a general approximation, our findings are not compatible with an assumption of impulsive decision-making and in line with previous reports that did not correct for reaction time delays [10]. This finding suggests that impulsiveness does not necessarily govern decisions in SCZ. We speculate that the decisions often occurred after just a few bead views because of impairments in stable belief-formation. Further research is needed to reconfirm our finding and to disentangle the underlying processes of increased time expenditure and decision making based on fewer bead views.

Finally, SCZ patients tended to overestimate the probabilities conveyed by the bead sequences presented in easy and difficult trials of the graded-estimates (GE) version (see Figure 3A), which is in line with previous reports [12, 13, 27, 55, 61]. Of note, these findings appeared not to be due to a global cognitive impairment, as SCZ patients were able to adapt their responses from easy to difficult and to ambiguous trials. In addition to assigning increased probabilities to viewed bead sequences, patients with SCZ showed increased disconfirmatory-evidence-scores (DES), whenever there was an unexpected change in bead color in a bead sequence after ≥2 beads of the same color (Figure 3B). This indicates that SCZ patients assign increased levels of significance to unexpected changes in bead color and struggle to form stable beliefs about the source jar as suggested by previous findings [9, 10, 12, 55, 57]. Of note, we employed the same

trial-sequences in both the DTD and the GE versions of the task, respectively. Hence, JTC was only measured when patients could limit data gathering by making a decision and stopping the current trial (DTD version of the task). However, when SCZ patients were presented with more bead views from the same trial-sequence in the GE version, they did not stick to their 'JTC choice'. Rather, each time unexpected changes in bead sequences occurred (e.g. a green bead after the three preceding beads were blue) they switched their probability ratings towards the opposite source jar (e.g. switching from rating a high probability for the predominantly blue after the first three blue beads to a medium probability for the source jar containing more green beads). We also observed changes of probability ratings following changes in bead sequences in MDD patients and HC participants. However, we only observed significantly increased DES and more pronounced changes in patients with SCZ. In sum, the GE version of the beads task therefore appears to be more suitable for detecting unstable belief-formation in patients with schizophrenia in addition to the JTC bias that can be obtained from the DTD version. Of interest in this regard, reduced DTD and increased JTC in SCZ were only observed in the $P_{80/20}$ condition, but not in the more difficult $P_{60/40}$ condition, although the very same trial sequences and trial difficulty levels were applied (see Supplemental Tables 5-6). SCZ patients therefore appeared as able as MDD and HC to adapt their performance to more difficult trial types and task conditions. However, future research is needed to clarify the role of probabilistic reasoning on beads task performance [61] and to rule out the contribution of sequence effects to this finding as our two conditions ($P_{80/20}$ / $P_{60/40}$) were presented in fixed order.

Limitations and Conclusions

As we had adopted a group-comparison design, participants within each group had been thoroughly matched with respect to their sociodemographic, neuropsychological, and clinical characteristics. Additionally, PANSS scores suggest that SCZ patients were mildly paranoid, and we did not investigate acutely ill patients nor did we assess JTC longitudinally across the disease course. Further, we investigated whether JTC was associated with current severity of delusions, not with their occurrence. Finally, our

sample size was only sufficiently powered to detect PDI differences across all groups and our sample was not representative of the psychoses continuum in the general population [60]. Future studies could usefully address these issues by investigating larger samples regarding PDI associations and test both delusional and non-delusional patients with SCZ longitudinally in naturalistic designs.

Specific advantages of our study are that we could demonstrate for the first time the important contributory roles of unstable beliefs and individual decision times to JTC, thereby challenging the notion that JTC reflects impulsive decision-making. As current psychotherapeutic and psychosocial interventions for patients with schizophrenia aim to modify impulsive decision-making, these findings have implications to inform further developments in psychotherapeutic treatment such as meta-cognitive training [29, 30]. Importantly, our findings further imply a discussion about the construct validity of different versions of the beads task. By employing a novel beads task design that further evaluates explicit probability estimates and disconfirmatory evidence responses we were able to observe that SCZ patients tended to form unstable beliefs. Additionally, SCZ patients needed more time than MDD or HC to make their decisions. As the JTC bias has been demonstrated to be modifiable [29, 62] this contribution to our understanding of its components offers the potential to inform the further development of related treatment options.

Acknowledgements and Funding

We would like to thank all participants for so generously giving their time. This work was supported by a research grant of the Deutsche Forschungsgemeinschaft (No. STR 1472/1-1), by a research grant of the Friedrich-Baur-Stiftung (No. FBS 59-17), and by a research grant of the Medical Faculty, University of Augsburg (No. 89000101) (all to WS). The funders had no role in the design and conduct of the study, the collection and management of the data as well as their respective analysis and the preparation or approval of the manuscript.

Conflict of interest statement

Wolfgang Strube has received a speaker's honorarium from Mag&More GmbH. Peter Falkai was honorary speaker for Janssen-Cilag, AstraZeneca, Eli Lilly, Bristol–Myers–Squibb, Lundbeck, Pfizer, Bayer Vital, SmithKline Beecham, Wyeth and Essex. During the last 5 years, but not presently, he was a member of the advisory boards of Janssen-Cilag, AstraZeneca, Eli Lilly and Lundbeck. Alkomiet Hasan has received a paid speakership from Janssen-Cilag, Otsuka and Lundbeck. He was a member of an advisory board of Roche, Janssen-Cilag, Otsuka and Lundbeck. Sven Bestmann and all other authors, Camelia Lucia Cimpianu, Miriam Ulbrich, Ömer Faruk Öztürk, Thomas Schneider-Axmann and Louise Marshall, declare no competing financial interests.

Author Contributions

Conceptualization: WS SB AH. Formal analysis: WS SB AH TSA. Funding acquisition: WS SB. Investigation: WS CC MU. Methodology: WS SB AH TSA. Project administration: WS AH. Software: WS SB LM. Supervision: WS SB AH. Visualization: WS SB LM. Writing – original draft: WS AH SB. Writing – review & editing: all authors.

Open Science Statement

The Matlab code for our experimental beads task instantiation can be downloaded at https://github.com/wstrube/Beadstask.

Study Registration Details

Trial name: "Behavioral investigation of the influence of impaired neurotransmission on perceptual and decision-making processes using computer-assisted mathematical model systems in people with schizophrenia and depression."

URL:

https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00019012

(webpage last accessed on 5th Juli 2021)

https://www.drks.de/drks_web/setLocale_EN.do

(webpage last accessed on 5th Juli 2021)

Registration Number: DRKS00019012

References

References

1. Dudley R, Taylor P, Wickham S, Hutton P. Psychosis, Delusions and the "Jumping to Conclusions" Reasoning Bias: A Systematic Review and Meta-analysis. *Schizophr Bull* May 2016;42(3):652-665.

2. Fine C, Gardner M, Craigie J, Gold I. Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cogn Neuropsychiatry* Jan 2007;12(1):46-77.

3. Garety PA, Freeman D. The past and future of delusions research: from the inexplicable to the treatable. *Br J Psychiatry* Nov 2013;203(5):327-333.

4. So SH, Siu NY, Wong HL, Chan W, Garety PA. 'Jumping to conclusions' data-gathering bias in psychosis and other psychiatric disorders - Two meta-analyses of comparisons between patients and healthy individuals. *Clin Psychol Rev* Jun 2016;46:151-167.

5. Garety PA, Freeman D. Cognitive approaches to delusions: a critical review of theories and evidence. *Br J Clin Psychol* Jun 1999;38 (Pt 2):113-154.

6. Woodward TS, Mizrahi R, Menon M, Christensen BK. Correspondences between theory of mind, jumping to conclusions, neuropsychological measures and the symptoms of schizophrenia. *Psychiatry Res* Dec 30 2009;170(2-3):119-123.

7. Huq SF, Garety PA, Hemsley DR. Probabilistic judgements in deluded and non-deluded subjects. *Q J Exp Psychol A* Nov 1988;40(4):801-812.

8. Hemsley DR, Garety PA. The formation of maintenance of delusions: a Bayesian analysis. *Br J Psychiatry* Jul 1986;149:51-56.

9. Garety PA, Hemsley DR, Wessely S. Reasoning in deluded schizophrenic and paranoid patients. Biases in performance on a probabilistic inference task. *The Journal of nervous and mental disease* Apr 1991;179(4):194-201.

10. Moritz S, Woodward TS. Jumping to conclusions in delusional and non-delusional schizophrenic patients. *Br J Clin Psychol* Jun 2005;44(Pt 2):193-207.

11. Evans SL, Averbeck BB, Furl N. Jumping to conclusions in schizophrenia. *Neuropsychiatr Dis Treat* 2015;11:1615-1624.

12. Adams RA, Napier G, Roiser JP, Mathys C, Gilleen J. Attractor-like dynamics in belief updating in schizophrenia. *Journal of Neuroscience* 2018:3163-3117.

13. Speechley WJ, Whitman JC, Woodward TS. The contribution of hypersalience to the "jumping to conclusions" bias associated with delusions in schizophrenia. *J Psychiatry Neurosci* Jan 2010;35(1):7-17.

14. Moritz S, Woodward TS, Lambert M. Under what circumstances do patients with schizophrenia jump to conclusions? A liberal acceptance account. *Br J Clin Psychol* Jun 2007;46(Pt 2):127-137.

15. Averbeck BB, Evans S, Chouhan V, Bristow E, Shergill SS. Probabilistic learning and inference in schizophrenia. *Schizophr Res* Apr 2011;127(1-3):115-122.

16. Menon M, Pomarol-Clotet E, McKenna PJ, McCarthy RA. Probabilistic reasoning in schizophrenia: a comparison of the performance of deluded and nondeluded schizophrenic patients and exploration of possible cognitive underpinnings. *Cogn Neuropsychiatry* Nov 2006;11(6):521-536.

17. Pfuhl G. A Bayesian perspective on delusions: Suggestions for modifying two reasoning tasks. *J Behav Ther Exp Psychiatry* Sep 2017;56:4-11.

18. Pytlik N, Soll D, Hesse K, et al. Problems in measuring the JTC-bias in patients with psychotic disorders with the fish task: a secondary analysis of a baseline assessment of a randomized controlled trial. *BMC Psychiatry* Nov 23 2020;20(1):554.

19. Garety PA, Joyce E, Jolley S, et al. Neuropsychological functioning and jumping to conclusions in delusions. *Schizophr Res* Nov 2013;150(2-3):570-574.

20. Lincoln TM, Ziegler M, Mehl S, Rief W. The jumping to conclusions bias in delusions: specificity and changeability. *J Abnorm Psychol* Feb 2010;119(1):40-49.
21. McLean BF, Mattiske JK, Balzan RP. Association of the jumping to conclusions and evidence integration biases with delusions in psychosis: a detailed meta-analysis. *Schizophrenia bulletin* 2017;43(2):344-354.
22. Peters ER, Thornton P, Siksou L, Linney Y, MacCabe JH. Specificity of the jump-to-conclusions bias in deluded patients. *British Journal of Clinical Psychology* 2008;47(2):239-244.
23. Falcone MA, Murray RM, O'Connor JA, Hockey LN, Gardner-Sood P, Di Forti M, Freeman D, Jolley S. Jumping to conclusions and the persistence of delusional beliefs in first episode psychosis. *Schizophr Res* Jul 2015;165(2-3):243-246.
24. Moritz S, Woodward TS, Hausmann D. Incautious reasoning as a pathogenetic factor for the development of psychotic symptoms in schizophrenia. *Schizophrenia Bulletin* 2005;32(2):327-331.
25. Bentham AM, McKay AP, Quemada I, Clare L, Eastwood N, McKenna PJ. Delusions in schizophrenia: a phenomenological and psychological exploration. *Cogn Neuropsychiatry* Nov 1 1996;1(4):289-304.
26. Ashinoff BK, Singletary NM, Baker SC, Horga G. Rethinking delusions: A selective review of delusion research through a computational lens. *Schizophr Res* Mar 3 2021.
27. Baker SC, Konova AB, Daw ND, Horga G. A distinct inferential mechanism for delusions in schizophrenia. *Brain* Jun 1 2019;142(6):1797-1812.
28. Sterzer P, Adams RA, Fletcher P, et al. The Predictive Coding Account of Psychosis. *Biol Psychiatry* May 25 2018.
29. Moritz S, Veckenstedt R, Bohn F, et al. Complementary group Metacognitive Training (MCT) reduces delusional ideation in schizophrenia. *Schizophr Res* Dec 2013;151(1-3):61-69.
30. Morrison AP, Renton JC, Dunn H, Williams S, Bentall RP. *Cognitive therapy for psychosis: A formulation-based approach.* New York, NY: Brunner-Routledge; 2004.
31. Tripoli G, Quattrone D, Ferraro L, et al. Jumping to conclusions, general intelligence, and psychosis liability: findings from the multi-centre EU-GEI case-control study. *Psychol Med* Apr 24 2020:1-11.
32. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* Sep 2011;21(9):655-679.
33. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 2007;39:175-191.
34. Strube W, Marshall L, Quattrocchi G, et al. Glutamatergic Contribution to Probabilistic Reasoning and Jumping to Conclusions in Schizophrenia: A Double-Blind, Randomized Experimental Trial. *Biol Psychiatry* Nov 1 2020;88(9):687-697.
35. Guy W, Bonato R. CGI: Clinical Global Impressions. *ECDEU assessment manual for psychopharmacology*: U.S. National Institute of Health; 1976:218-222.
36. Hall RC. Global assessment of functioning: a modified scale. *Psychosomatics* 1995;36(3):267-275.
37. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987;13(2):261-276.
38. Wallwork RS, Fortgang R, Hashimoto R, Weinberger DR, Dickinson D. Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale for schizophrenia. *Schizophr Res* May 2012;137(1-3):246-250.
39. Beck AT, Steer R. Manual for the revised Beck depression inventory. *San Antonio, TX: Psychological Corporation* 1987.
40. Hamilton M. A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry* 1960;23(1):56.

41. Peters E, Joseph S, Day S, Garety P. Measuring delusional ideation: the 21-item Peters et al. Delusions Inventory (PDI). *Schizophrenia bulletin* 2004;30(4):1005.

42. Brickenkamp R, Zillmer E. *The d2 test of attention*: Hogrefe & Huber Seattle, WA; 1998.

43. Reitan R. TMT, Trail Making Test A & B. *South Tucson, AR: Reitan Neuropsychology Laboratory* 1992.

44. Schmidt K-H, Metzler P. *Wortschatztest: WST*: Beltz; 1992.

45. UCL Labratory of Neurobiology FIL, and Institute of Cognitive Neuroscience. Cogent Graphics (LON) and Cogent 2000 (FIL and ICN). Graphics toolboxes for Matlab. Available at: http://www.vislab.ucl.ac.uk/cogent.php.

46. Phillips LD, Edwards W. Conservatism in a simple probability inference task. *J Exp Psychol* Sep 1966;72(3):346-354.

47. Cisek P, Puskas GA, El-Murr S. Decisions in changing conditions: the urgency-gating model. *J Neurosci* Sep 16 2009;29(37):11560-11571.

48. Ormrod J, Shaftoe D, Cavanagh K, Freeston M, Turkington D, Price J, Dudley R. A pilot study exploring the contribution of working memory to "jumping to conclusions" in people with first episode psychosis. *Cogn Neuropsychiatry* 2012;17(2):97-114.

49. Mehta CR. The exact analysis of contingency tables in medical research. *Stat Methods Med Res* 1994;3(2):135-156.

50. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in sport sciences* 2014;1(21):19-25.

51. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophr Res* Nov 15 2005;79(2-3):231-238.

52. Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton Depression Rating Scale. *J Affect Disord* Sep 5 2013;150(2):384-388.

53. Garety PA, Freeman D, Jolley S, Dunn G, Bebbington PE, Fowler DG, Kuipers E, Dudley R. Reasoning, emotions, and delusional conviction in psychosis. *Journal of abnormal psychology* 2005;114(3):373.

54. Peters ER, Joseph SA, Garety PA. Measurement of delusional ideation in the normal population: introducing the PDI (Peters et al. Delusions Inventory). *Schizophrenia bulletin* 1999;25(3):553-576.

55. Jardri R, Duverne S, Litvinova AS, Deneve S. Experimental evidence for circular inference in schizophrenia. *Nat Commun* Jan 31 2017;8:14218.

56. Heinz A, Schlagenhauf F. Dopaminergic dysfunction in schizophrenia: salience attribution revisited. *Schizophr Bull* May 2010;36(3):472-485.

57. Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry* Jan 2003;160(1):13-23.

58. Langdon R, Ward PB, Coltheart M. Reasoning anomalies associated with delusions in schizophrenia. *Schizophr Bull* Mar 2010;36(2):321-330.

59. Peters E, Garety P. Cognitive functioning in delusions: a longitudinal analysis. *Behav Res Ther* Apr 2006;44(4):481-514.

60. Ross RM, McKay R, Coltheart M, Langdon R. Jumping to Conclusions About the Beads Task? A Meta-analysis of Delusional Ideation and Data-Gathering. *Schizophr Bull* Sep 2015;41(5):1183-1191.

61. Klein HS, Pinkham AE. Examining reasoning biases in schizophrenia using a modified "Jumping to Conclusions" probabilistic reasoning task. *Psychiatry Res* Dec 2018;270:180-186.

62. So SH, Freeman D, Dunn G, Kapur S, Kuipers E, Bebbington P, Fowler D, Garety PA. Jumping to conclusions, a lack of belief flexibility and delusional conviction in psychosis: a longitudinal investigation of the structure, frequency, and relatedness of reasoning biases. *J Abnorm Psychol* Feb 2012;121(1):129-139.