Delusional Beliefs: Exploring Associated Cognitive and Emotion Factors

Daisy Fitzpatrick

A thesis submitted in partial fulfilment of the requirements for the award of Doctorate in Clinical Psychology at the University of Sheffield

Clinical Psychology Unit, Department of Psychology

Faculty of Science

The University of Sheffield

May 2024

Declaration

I declare that this work has not been submitted for any other degree at the University of Sheffield or any other institution. This thesis is my own original work and all other sources have been referenced accordingly.

Structure and Word Count

Section 1: Literature Review

Excluding references and tables: 7981

Including references and tables: 13925

Section 2: Empirical Study

Excluding references and tables: 7996

Including references and tables: 10700

Total Word Count:

Excluding references and tables: 15977

Including references and tables: 24625

Acknowledgements

First and foremost, I would like to thank my research supervisor, Professor Richard Bentall, for his encouragement, guidance, and support throughout the entirety of the project. I have felt inspired by his passion for conducting research in this area and the knowledge shared has been invaluable.

I would like to thank all the participants who kindly offered their time to complete this study, this project would not have been possible without your generosity. I would also like to express thanks for the advice and support provided by experts with experience and clinicians working in the field, who helped shape the project.

I would like to thank my fellow trainee clinical psychologist, Jess Twigg, for both her role in the practical elements of the project but also her support as a close friend. Thank you for collaborating on data collection with me and acting as a second reviewer for the systematic review. Also, thank you for your unwavering support throughout the entirety of my DClin journey. I will forever cherish the friendship we have formed and I cannot imagine having completed this journey without you!

I would finally like to thank my amazing family, friends and colleagues, whose support has been instrumental in helping me achieve my dreams of becoming a clinical psychologist. A special thank you to my partner Jamie who has been an unwavering source of support.

Lay Summary

Delusional beliefs are described as strongly held false beliefs which are not open to change. There are different sub-types of delusional beliefs with persecutory, reference and grandiose beliefs most commonly experienced. There is currently no agreement within the literature on why people experience these beliefs. Both sections of this thesis aimed to better understand this.

Section 1 specifically considered grandiose delusions (GD), which are described as false beliefs about a special identity or inflated sense of power, worth or knowledge. A systematic literature review was carried out to identify which emotional factors (i.e., anxiety, self-esteem, depression) are present in adults who experience GD.

Three online databases were searched and 21 studies were found. Studies considered a range of emotions and used different measures to assess them and also GD, making comparisons difficult. Depression, self-esteem and anxiety were most frequently looked at. Overall, the findings were mixed. Findings suggest a link between GD and lower depression and no link with increased anxiety. Findings for self-esteem were more conflicting. Confidence in review findings was low due to the varied methodological quality of the studies. The review highlighted a lack of research on GD.

Section 2 considered a different explanation for delusional beliefs. It has been suggested that the way people judge how certain they are about their beliefs may be different in the case of patients with delusions. When this has been examined in previous research, studies have usually considered neutral beliefs such as general knowledge quiz answers, rather than threat-related beliefs. This may be important as threat is the most common content of delusional beliefs.

A total of 66 participants were recruited via social media and NHS services to three equal groups: 1) participants experiencing delusional beliefs, 2) those experiencing

depression and/or anxiety but not delusional beliefs, and 3) individuals not experiencing delusional beliefs or depression and/or anxiety. Participants rated how certain they were when reporting their attitudes towards neutral items (i.e. rugby) and future events. Participants were also asked 24 multiple-choice general knowledge questions, with 12 being neutral and 12 threat-related content (e.g., about crime names/definitions). Participants rated their certainty in answers and their answer response time (RT) was recorded.

Participants were more certain about correct answers and this did not differ for individuals experiencing delusions. Most individuals were also quicker to respond with answers they were more certain about, but this was not true for individuals experiencing delusional beliefs whose RTs were not associated with their certainty. Overall, patients experiencing delusional beliefs do not appear to differ in self-rated certainty but do differ in their related RTs (longer RT). Questions being neutral or threat-related did not significantly impact certainty judgements in any group. Future research is needed to replicate findings in larger clinical samples.

Table of Contents

Section 1: Literature Review

Emotion Factors Associated with Grandiose Delusions: A Systematic Review

Abstract

Objectives

Grandiose Delusions (GD) are described as false beliefs about a special identity or inflated sense of power, worth or knowledge. Theoretical understandings of psychological factors associated with GD remain inconclusive. The review aimed to explore the current literature on GD, focusing specifically on associated emotion factors.

Methods

Following a published protocol (PROSPERO ID: CRD42024505574), a systematic review was completed in January 2024 across PsycINFO, MEDLINE and Scopus to identify quantitative studies reporting on emotion factors related to GD, in adults. Methodological quality and risk of bias for included studies was evaluated using an appraisal tool and a narrative synthesis was carried out.

Results

Twenty-one studies were included in the final narrative synthesis. Large methodological heterogeneity was present within studies and methodological quality varied. Eight emotion factors were identified with depression, anxiety and self-esteem being the most commonly explored. Relationships between emotion factors and GD were inconsistent across studies, particularly for self-esteem, however, there appeared to be some support for a link with lower levels of depression and no link with increased anxiety.

Conclusions

The review highlighted a scarcity of research exploring emotion factors associated with GD and limitations with study designs prevented the exploration of causality. Given the heterogeneity and varying methodological quality caution is required within interpretation. Better-quality research utilising experimental and longitudinal study designs is needed to understand the role of emotion in the formation and maintenance of GD. Greater

understanding would support with identifying at-risk populations and the development of interventions for use within clinical practice.

Keywords: Psychosis, grandiose delusions, delusions of grandeur, emotion, affect

Practitioner Points

- There is a dearth of research exploring emotion factors associated with grandiose delusions.
- Lower depression may be associated with grandiose delusions but further research is needed.
- No studies found an association between grandiose delusions and higher anxiety.
- Associations of depression, self-esteem and anxiety with grandiose delusions were the emotion factors most commonly researched within studies.

Introduction

Delusional beliefs are commonly reported in the context of psychotic disorders including schizophrenia, but they can also occur in individuals with a range of psychiatric diagnoses (Picardi et al., 2018). The Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-5; American Psychiatric Association [APA], 2013) describes these beliefs as different from other strongly held beliefs in the certainty with which they are held despite reasonable contrary evidence, with literature often describing delusional beliefs as not amenable to change (Kiran & Chaudhury, 2009; Moritz et al., 2015). A recent review of the global prevalence rates of different types of delusional beliefs identified persecutory, reference, and grandiose beliefs as the three most commonly experienced, being present in 64%, 38.7% and 28.5% of participants respectively (Collin et al., 2023). There were few differences in these prevalence rates between parts of the world with very different economic and cultural circumstances.

Grandiose Delusions (GD), sometimes termed delusions of grandeur, or expansive delusions (Corsini, 2001), are described as false beliefs about a special identity or inflated sense of power, worth or knowledge maintained despite contradictory evidence (APA, 2013). GD are estimated to occur in around 50% of patients with a schizophrenia diagnosis and around 66% of patients with a bipolar disorder diagnosis (Knowles et al., 2011). Literature has found GD are experienced as meaningful by individuals, helping them to make sense of difficult situations or by providing a sense of purpose and self-identity. Despite this, a wide range of negative impacts including social, emotional, sexual, physical, and occupational harms have also been linked with their presence (Isham et al., 2021).

Theories of (Grandiose) Delusions

The mechanisms behind the formation and maintenance of delusional beliefs are still somewhat unclear, with no one theory being widely accepted as a singular explanation of

delusional beliefs (Connors & Halligan, 2020). Different theories have been proposed to explain the formation and maintenance of delusional beliefs including, social cognitive theories emphasising the role of past traumatic experiences (Bailey et al., 2018); cognitive theories such as maladaptive belief evaluation (Langdon & Coltheart, 2000), and those centred around affect. Researchers have commented on the array of evidence to support the role of emotions in both the cause and maintenance of delusions, suggesting that delusions may be a reflection of internal emotional states (Freeman & Garety, 2003).

Although argued to have received less focus within the theoretical and empirical literature than their persecutory counterparts, several theories of GD have been suggested (Knowles et al., 2011). Two main emotion-related theoretical accounts proposed are: delusions-as-defence (DAD; Neale, 1988) and "emotion-consistent" accounts (Freeman & Garety, 2003; Smith et al., 2005). DAD, first suggested as an explanation for persecutory delusions, centres on the notion delusions emerge as a protective mechanism around negative emotional states. Literature has suggested GD may develop to counteract feelings of powerlessness and loneliness (Beck & Rector, 2005), with a suggestion of the emergence of GD as a result of incongruence between desired and actual self and to protect from low self-esteem (Neale, 1988). Contrastingly, emotion-consistent accounts propose that delusions develop to reflect internal emotion states, with GD emerging due to positive emotions and representing aspects of the self held with higher regard (Freeman & Garety, 2003; Smith et al., 2005).

There are still inconsistencies within the literature about which explanation provides the best fit and past reviews have attempted to understand the processes involved in the formation and maintenance of GD. Knowles et al. (2011) conducted a review exploring the existing cognitive and emotion literature on GD, examining support for both emotion-consistent and DAD theories. The review included papers with both adult and child samples

greater than 50. They comment on some psychological factors associated with GD such as direct and protective roles of emotion and self-esteem. They suggested that emotion-consistent accounts had the greatest support but did not rule out DAD theories. As a result of their review, they proposed a new model encompassing both the theories mentioned above. They propose a role of precipitating events influencing an internal state change, which through a search for meaning leads to an appraisal. This appraisal then prompts an initial grandiose thought with both of these then contributing to GD. They highlight the potential influence of rumination, life events, culture, cognitive bias, unstable fluctuating self-esteem, and persecutory delusions within this relationship.

Aims and Objectives

Despite attempts within the literature to understand the formation and maintenance of delusional beliefs, definitive psychological factors associated with GD have yet to be established, with no recent systematic review having been carried out looking at GD and their relationship with a broader range of emotion factors.

This systematic review aimed to synthesise the current literature regarding emotional factors (i.e., self-esteem, anxiety, low mood, etc.) related to GD in adult populations to support theoretical understanding and subsequently clinical practice.

Method

The systematic review was conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Page et al., 2021; see Appendix A) and pre-registered on PROSPERO in January 2024 (accessible via the following link: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42024505574).

Search Strategy

Title, abstract and keyword searches were conducted on 31.01.2024 across three databases PsycINFO, MEDLINE and Scopus.

No limits were applied to the search regarding date or language at the search stage as the review was interested in a comprehensive overview of all papers and wished to maximise search results. No study registers, regulatory databases or other online repositories were searched, and no individuals or organisations were contacted to identify studies.

Search terms were generated from key terms considering those used in a previous similar review (Knowles et al., 2011), alongside guidance from a research supervisor and the University of Sheffield library service. The search strategy used variations of the search terms 'delusions', 'psychosis', or 'schizophrenia' in combination with variations in 'grandiose' or 'grandeur' and variations of 'cognitive', 'emotion' or 'psychological', accounting for synonyms, variations, and truncations. Groups of search terms were searched individually before being combined with the Boolean operator 'AND' using the OVID tool. Each group of terms were first searched individually and then combined with the 'AND' Boolean operator to generate the final search. This resulted in the following search string which was mapped and applied across databases: (delusion* OR delud* OR psychosis OR schizophreni* OR psychotic) AND (grandeur* OR grandi* OR Expansive) AND (cogni* OR emoti* OR psychol*).

Supplementary Search

Although systematic reviews and meta-analyses were not included in the current review, those identified through the search and felt to be appropriate were manually searched for relevant papers to be included (Collin et al., 2023; Grbic, 2013). Additionally, manual forward (via Google Scholar and OVID) and backwards (reviewing reference lists) citation searching were conducted on all included papers to increase confidence that all relevant papers were identified. The decision was made to exclude grey literature from the search.

Although searching and including grey literature would allow for a more comprehensive overview of the literature, it is often non-peer-reviewed and can be of lower quality (Benzies et al., 2006).

Inclusion/Exclusion Criteria

Studies were screened against pre-determined inclusion and exclusion criteria using the Population, Intervention, Comparator, Outcome and Study design framework (PICOS) in line with PRISMA guidance. See Table 1.

Quantitative studies were included if they reported on the relationship of GD with at least one emotion factor in adult populations (18 years or over). For the purpose of the review emotional factors are defined as affective elements that influence an individual's thoughts, attitudes, feelings, and behaviours (i.e., self-esteem, anxiety, low mood, etc.). Studies where GDs were indistinguishable from other delusion types in the reporting of relationships, or where GDs were linked with neurological conditions, delirium, or occurred only in the context of personality disorders were excluded. Only studies published in peer-reviewed journals, and with a transcript available in English were included.

Two changes were made following the original registered protocol. Originally the review question focussed on psychological factors (both cognitive and emotional factors) associated with GD. Due to time constraints and a large number of papers identified meeting the criteria during the piloting of study selection, the decision was made to focus only on emotional factors associated with GD. Secondly, initial inclusion criteria included a psychiatric diagnosis relating to psychosis for the population sample, however, a more inclusive approach was adopted and the protocol was amended to remove the requirement, given some identified papers included individuals with a variety of psychiatric diagnoses and those with GD in the absence of a psychosis-related diagnosis.

Table 1

Inclusion and Exclusion Criteria Using the PICOS Framework

PICOS domain	Inclusion	Exclusion
Population	• Adults aged 18 years old and over who are experiencing GD. • No other restrictions were applied in relation to setting, location, gender.	• Individuals aged under 18 years old, including samples where adults cannot be distinguished from under 18-year-old populations. • Individuals with neurological conditions or delirium that may be linked with the presence of GD. Individuals with GD in the context of personality disorders will be excluded.
Intervention/ exposure	• Experience of GD with measure allowing distinction from others type of delusions.	• Studied where experience of GD cannot be distinguished from other types of delusions.
Comparator	• Studies with and without comparison controls (adults aged 18 years old and above who do not experience GD).	• No comparison group is required.
Outcomes	• Studies specifically reporting the relationship between GD and at least one emotion factor. Both validated and non-validated measures of delusions and factors were eligible.	• Studies where a relationship between the variables and GD specifically are not reported.
Study design	• Quantitative, peer reviewed, empirical studies, available in the English language. • No restrictions on date or location.	• Studies not available in English. • Qualitative studies, case studies or reports, non-empirical literature (e.g., book chapters, review articles, anecdotal papers, conference papers), unpublished literature, and grey literature.

Note. GD = grandiose delusions

Selection

Papers identified through the initial search were exported to Zotero (reference manager) and duplicates were identified through the automation tool, with manual review and removal of duplicates. No other automation tools were used. Papers were then uploaded to Covidance (systematic review tool) due to its specific user interface for screening and its ability to randomly allocate a percentage of papers to a second reviewer at the various stages (title/abstract screen, full-text review, data extraction, and quality assessment).

The author (D.F) screened titles and abstracts of all papers identified within the initial search against the inclusion/exclusion criteria. Following this full-texts were gathered for papers identified through the initial screening and a full-text review was carried out by the author (D.F).

A second reviewer (J.T) independently reviewed a selection of papers (randomly assigned) against inclusion/exclusion criteria at the different review stages, as this can increase reliability and minimise error and selection bias within the review process (Stoll et al). Both reviewers were blinded to the other reviewer's rating and conflicts were to be agreed through discussion. This included 10% of title abstract screening (n=104), and 10% of full-text review for inclusion (n=15). For papers where there was a disagreement, the decision was discussed between reviewers. A 100% consensus was reached between reviewers following discussions.

A PRISMA flowchart diagram was created to detail each stage of the search and selection process (see results, Figure 1).

Data Extraction

Data was extracted from all papers identified for inclusion in the review using pre-established extraction categories. A second reviewer (J.T) independently extracted 100% of

the papers (n=21), with extracted information from both reviewers being collated into a study characteristics table. One paper did not report the exact p number, referring to it as non-significant. Due to time constraints, the author was not contacted to obtain this. Extraction categories included: author(s) and year, country, design, sample characteristics and setting, measures of GD and emotion factors, and key statistical findings relating to the relationship between these.

Quality Assessment

Included studies were assessed for methodological quality and risk of bias using an adapted version of the National Heart, Lung, and Blood Institute (NHLBI, 2013) quality assessment for observational studies tool (Appendix B).

When deciding on an appropriate tool the Newcastle-Ottawa Scale (NOS; Wells et al., 2021) was explored as a potential tool due to being identified as a commonly used tool for observational studies within reviews (Ma et al., 2020). However, cohort and case-control study design versions only are available. Subsequently, the NHLB was selected due to its applicability to multiple observational study designs including cohort and cross-sectional studies which formed the majority of included papers. Having a single quality appraisal tool appropriate for all included study designs was felt to be beneficial to support the comparison of methodological quality across all included studies within the review. Additionally, the NHLBI quality assessment tool uses a checklist approach which is suggested to demonstrate a better overview of study quality compared to approaches using a numerical rating (used in NOS) where detail can be lost regarding specific sub-elements that may be rated differently depending on the audience (Boland et al., 2017)

The NHLBI quality assessment tool is a 14-item tool assessing methodological quality relating to a study's aim, design, sampling (population, sample size justification/considerations), recruitment, identification and adjustment of confounding

variables, measure of IVs and DVs, and follow-up and attrition. Studies were rated against a checklist using the following: Yes (item adequately addressed), No (item not adequately addressed), P (partial: item partially addressed), NR (not reported), and NA (not applicable).

To increase reliability and to reduce potential performance bias (Gold et al., 2012), quality appraisal was carried out independently by both the lead author (D.F) and second reviewer (J.T.) for 100% of the papers. Both reviewers initially reviewed five papers before discussing ratings and reaching a shared consensus. This was repeated for every five papers.

Data Synthesis

In line with Economic and Social Research Council recommendations (Popay et al., 2006) a narrative synthesis was conducted. A meta-analysis was considered for instances where the search identified five or more papers exploring an emotion using the same measure. However, given the degree of clinical heterogeneity between studies (samples, research questions), the potential overlap of participant samples within multiple studies (Fowler et al., 2006; Garety et al., 2013; Smith et al., 2006), study quality issues, and recommendations regarding conducting meta-analyses (Lensen, 2023), through discussions as a research team it was decided a meta-analytic comparison was not appropriate.

Results

Search and Selection

Initial searches identified 1566 papers (PsycINFO= 399, MEDLINE= 377, Scopus= 790). Following the removal of 530 duplicate records, 1036 papers entered the title and abstract screening stage with 892 being excluded due to not meeting the inclusion criteria. Full texts for the remaining 144 papers were reviewed with 128 papers being excluded due to: no emotion comparison being distinguishable (n=80); incorrect study design (n=16); GD not distinguishable from other types (n=14); full-text not available in English (n=13); did not explore GD (n=3); non-peer-reviewed (n=1); and sample included participants <18 years old

(n=1). Some papers had multiple exclusion criteria applicable but only the first identified reason was noted and reported.

The remaining 16 papers were assessed by both reviewers as meeting the criteria for inclusion in the review. Further searching (of identified reviews and forward and backwards searching) resulted in five additional papers being identified, resulting in a total of 21 papers being included in the final narrative synthesis. Both reviewers reached 100% consensus on all 21 papers. See Figure 1 for the PRISMA flow diagram.

During the review process, one study was excluded due to participants aged 16 years old being included within the sample (Harder, 2006). Although the study was excluded for violating the inclusion/exclusion criteria, given some of the sample was above 18 years old it may have provided beneficial information for the review.

Study Characteristics

Details of study characteristics and key findings for all 21 included papers are presented in Table 2.

Figure 1

PRISMA Flow Chart Diagram Adapted from Page et al. (2021)

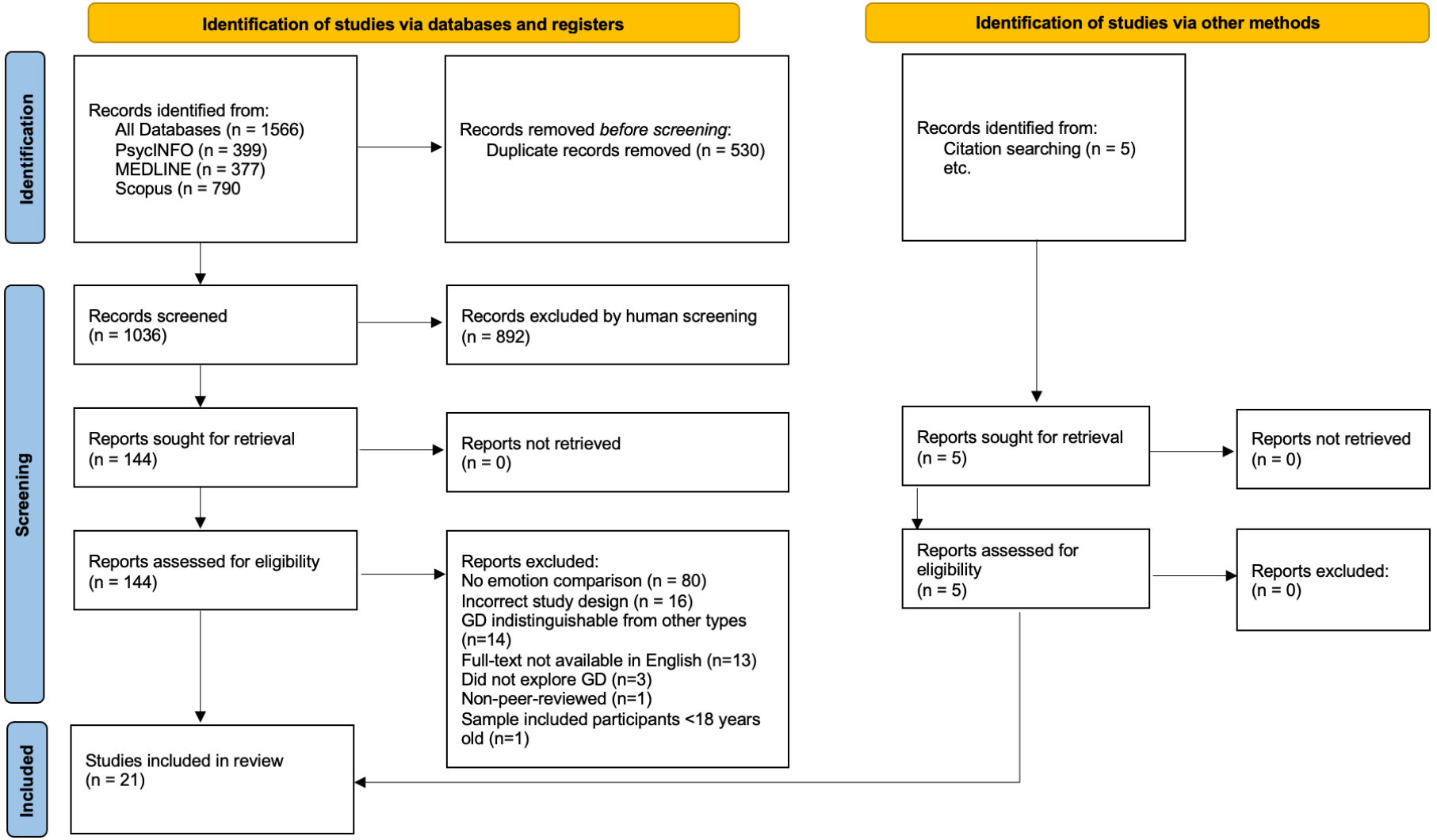


Table 2

Summary of Main Characteristics, GD and Emotion Measures, and Findings of Included Studies (Ordered Alphabetically by First Author Name)

Author(s) and year/ country	Design	Sample characteristics and setting	Measure(s) of GD	Emotion factor(s) and measure	Key findings
Appelbaum et al. (1999) USA	CS	Comparison group: no GD; Sample size: 328 (GD n=141); Setting: INP; Age range 18-40; Gender: NR; Ethnicity: NR; Diagnosis: NR (mixed for wider sample).	Researcher determined using 17 questions drawn primarily from the Diagnostic Interview Schedule based on DSM-III-R definition of a delusion. Where multiple type delusional beliefs were present participants self-selected the one most important to them or it was interviewer selected. MAMDAS then used.	Negative affect: (unhappy, frightened, anxious, or angry): MMADS.	Significant difference in negative affect between subjects with and without GD, being lower in individuals experiencing GD. ($p<.001$, separate one-way ANOVA, df=1, 315). No F value reported.
Ben-Zeev et al. (2012) USA	CS	Comparison group: none; Sample size: 130; Setting: Community, Mean age (SD): 46.2 (11.24); Gender: 59%	Self-report questionnaire: Single item for Grandiosity: "You had	Self-esteem: SERS-SF (negative subscale score)	With emotion factors as a predictor of GD using multilevel modelling, at the person level:

		Male; Ethnicity: 59% white, 15% African–American, 14% Hispanic, and 12% other ethnicities; Diagnosis: SZ or SAZ; mixed.	special powers to do something nobody else can do?" from PSYRATS.	Anxiety: Single question: “How anxious do you feel right now?” Sadness: Single question: ‘‘How sad do you feel right now?”	Lower self-esteem positively and significantly predicted the occurrence of GD (r=0.04 (SE=0.02), OR=1.04, 95% CI [1.01, 1.08], $p < .05$). Anxiety did not significantly predict the occurrence of GD (r=.09 (SE=0.08), OR=1.09 95% CI [0.93, 1.28], $p > .05$). Sadness did not significantly predict the occurrence of GD (r= -.14 (SE= 0.09), OR= 0.87, 95% CI [0.73, 1.04]) $p > .05$.
Bortolon et al. (2019) France	CS	Comparison group: none; Sample size: 115; Setting: INPT, OP; Mean age (SD): 36.91 (9.98); Gender: Male n=88 (76.5%); Female =27 (23.5%); Ethnicity: NR; Diagnosis: SZ.	PANSS (item 6)	Depression: CDSS	No significant correlations between GD and depression were found.
Boyden et al. (2015) UK	CS	Comparison group: depression diagnosis (DSM-IV); Sample size: clinical (GD) =18; Control (depressed) =14; Setting: Community -IAPT; Mean ages (SD): GD: 43.4 (9.1); Control: 43.5 (13); Gender:	Clinician confirmed; PDI-21	Depression: HADS Anxiety: HADS	Sig difference between GD and depressed control on both: Depression: Lower levels of depression in GD group compared to depressed control,

		GD = 56% Female; Control = 64% Female; Ethnicity: GD: All White British apart from 1(African Caribbean); Control: all white British Diagnosis DSM-IV: Clinical (GD): BD (n = 14), SZ (n = 2) or SAZ (n = 2); Control: Primary diagnosis of depression.			t(29.21) = 5.15 p<.0001, two-tailed test. Anxiety: Lower anxiety in GD group compared to depressed controls, t(29.09) = 4.28, p<.0001, two-tailed test.
de Portugal et al. (2009) Spain	CS	Comparison groups: other DD types; Sample size: Total=86 GD= 4; Setting: Community mental health centres; Mean age (SD): Total: 54 (14.4); GD: 56 (24.3); Gender: Total: 38.4% (n=33) Male, 61.6% (n=53) Female; GD: 2.3% Male, 2.3% Female (both n=2); Ethnicity: NR Diagnosis: 100% DD; 64% had premorbid PD.	SCID-I CV Clinician assigned to one of seven DD DSM-IV types (persecutory, jealous, somatic, erotomaniac, grandiose, mixed, and otherwise not specified).	Depression: MADRS	Scores of depressive symptoms were significantly lower amongst grandiose DD types than in the remaining types, p = .048 (Mann Whitney U). There was no significant difference in presence of depression in GD DD vs other types p = .118 (Mann Whitney U).
Fowler et al. (2006) UK	CS	Comparison group: non-clinical (students); Sample size: Clinical (psychosis): 252 (17% GD); Control: 754; Setting: INP, OP; Mean age (SD): Clinical: 38 (10); Control: 23.6	PANSS- 1+ positive PANSS rating (4+) PDI-40	Self-esteem: RSES	Non-clinical sample: A standard multiple regression with GD as a dependant and depression, paranoia, the RSES and the BCSS as

		(6.5); Gender: Clinical: 72% Male, 28% Female Control: 65% Female, 35% Male; Ethnicity: Clinical: White (n=187), Black-African (n=19), Black-Caribbean (n=19), Indian (n=5), other (n=19). Control: 70% were White, 17% Asian, 2% Afro-Caribbean, 1% African and 10% of other ethnic origin; Diagnosis: Clinical: SZ (n=217), SAZ (n=35), and DD (n=3); Control: "levels of depression, anxiety, paranoia and grandiosity were as expected for non-clinical group)"; PRP Trial.			explanatory was significant (F = 40.3; p < .001). RSES had weak trends, suggesting a small degree of unique association with GD (bivariate r = -.2; sr^2 = -.08; t = -2.4; p= .014). Clinical Sample: no comparison reported.
Garety et al. (2013) UK	CS	Comparison group: Persecutory delusions; Sample size: 280, GD only = 39, PD only = 134; Setting: INP, OP; Mean age (SD): GD: 39.3 (10.5) PD: 37.7 (11.1); Gender: GD: Male n=31/Female n=8; PD: Male n=91/Female n=43; Ethnicity: 72% White-British, 8% Black-Caribbean, 9% Black-African, and 11% other	Clinical diagnostic interview, SAPS item 11 score of between 2 (mild) and 5 (severe).	Depression: BDI-II; Self-esteem: RSES Anxiety: BAI	Persecutory vs GD GD were predicted by less negative self-evaluations, and lower anxiety and depression, along with higher positive self and positive other evaluations. Anxiety in GD found to be significantly lower compared to persecutory delusions group (OR=0.92 p = .003, CI= 0.86,-0.97)

		ethnic backgrounds; Diagnosis: SZ n=257 (85%) , SAZ n=40 (13%), and DD n=4 (2%); RPR Trial.			Depression in GD found to be significantly lower compared to persecutory delusions (OR =0.91 p< .001, CI=0.86-0.96) Self-esteem in GD found to be significantly higher compared to persecutory delusions (OR=0.82, p < .0001, CI=0.73-0.91).
Hartley et al. (2012) UK	CS	Comparison: none; Sample size: Total =229, GD = 146; Setting: NR, individuals were "receiving treatment"; Age: 37.9 (9.48) Gender: Male n= 196 and Female n= 33; Ethnicity: n= 93 (84.3%) of participants classed themselves as 'white' in ethnicity; Diagnosis: DSM-IV SZ, SZF, SAZ and DSM-IV diagnosis of drug and/or alcohol dependence or abuse; MIDAS trial.	Presence: PANSS Symptom severity: PSYRATS Symptom content: coded using symptom summary sheets.	Anxiety: PANSS Depression: PANSS	Depression: significantly lower depression with GD presence [t(206) = 1.97, p = .050] Anxiety: significantly lower anxiety levels were evident with GD presence [t(206) = 2.02, p = .045].
Hayashi et al. (2021) Japan	CS	Comparison group: none; Sample size: 108; Setting: INP =83 (77%); OP =25 (23%); Age: 51.8 (13.9) (range: 25-79); Gender: Male n=60;	Semi-structured interview, DOAI Section A. Delusion themes	Negative affect: DOAI; Depressed factor: PANSS depressed factor (anxiety, guilt feelings	GD were associated with lower depression and negative effect. GD and negative affective response (anxiety & tension, unpleasantness,

		Female n=48; Ethnicity: NR; Diagnosis: DSM-IV SZ, SAZ with persisting delusions.	PANSS	depression)	excitement and anger) r = -.421, p < .01 GD and PANSS depressed factor (anxiety, guilt feelings depression) r = -.209, p < .05.
Jacobsen et al. (2019) UK	CS	Comparison group: persecutory delusions; Sample size: Total = 41; GD = 11; PD = 30; Setting: INP; Mean age (SD): GD: 30 (24.54); PD: 33 (19.65); Gender: GD: Male n= 6 (55%); PD: Male n= 22 (73%) Ethnicity: GD: White 1 (9%), Asian 1 (9%), Black 5 (46%), Mixed Race 3 (27%), other 1 (9%); PD: White 9 (30%), Asian 3 (10%), Black 16 (53%), Mixed Race 2 (7%), other 0 (0%); Diagnosis: GD: SZ n= 8 (73%) MD n= 3 (27%); PD: SZ n=23 (77%) MD n= 7 (23%).	Clinician rating (2 clinicians) based on description of the delusion given by the participant; Self-ratings of psychotic symptoms.	Depression: DASS-21	(Mann Whitney U test) Depression significantly lower levels in GD compared to persecutory delusions Depression, U = 98.5, p = .05; No significant difference for Anxiety or Stress. Anxiety, U = 112, p= .118; Stress, U = 110, p= .105.
Moritz et al. (2010) Germany	Cross sectional/ Cohort	Comparison group: HC; Sample size: Study 1 (Baseline) Clinical = 58; HC = 44; Study 2 (4 weeks reassessed subsample) Clinical= 45, HC=24;	PANSS	Self-esteem: RSES	GD were modestly but significantly correlated with self-esteem (r = .26, p < .05). Comparison to control for GD

		Setting: INP, OP; Mean age: Clinical: 33.62 (11.95), HC: 31.91 (10.35); Gender: Clinical: Male n=41 Female n=17; HC: Male n=28, Female n=16; Ethnicity: NR; Diagnosis: Clinical: all SZ, n=14 Depression or Dysthymia; HC: No current or past psychiatric disorders.			not reported.
Schennach et al. (2015) Germany	Naturalistic/ Cohort	Comparison group: none; Sample size: 278; Setting: INP; Mean age: 34.77 (11.07); Gender: 163 Males, 115 Females; Ethnicity: NR; Diagnosis: DSM-IV 84% SZ, 12% SAZ and 4% brief psychotic disorder and in n= 11 comorbid depressive episode.	PANSS	Depression: CDSS	Significant negative relationship between GD and depression -patients with GD were less likely to suffer from depressive symptoms ($P < .01$). Strongest negative correlation was found between the PANSS item GD and the CDSS items: depressed mood, hopelessness, self-depreciation, morning depression and observed depression. Correlation: GD loaded negatively on Depression factor (DSS items: on depression, hopelessness and suicidality) = loading strength -0.45 (moderate in strength).

Schlier et al. (2015) Germany	CS	Comparison group: No Psychotic Disorder (NPD); Sample size: Total= 1046; Psychotic Disorder (Clinical)= 112; NPD = 934; Setting: NR for clinical sample/Community GP; Mean age: Clinical: 42.6 (12.9); NPD: 26.1 (9.0); Gender: Clinical: Male 58.1%, Female 41.9%; NPD: Male 31.7%, Female 68.3%; Ethnicity: NR; Diagnosis: SZ = 75.9%; SAZ = 16.1%; Brief psychotic disorder = 2.7% ; DD = 3.6%; Affective disorder with psychotic symptoms = 1.8%.	PANSS; PC	Depression: BDI-II; CES-D	When other CAPE factors were controlled for, they found no significant relationship (correlation) of GD with depression for both the BDI (n=222) = -.08; CES-D (n=207) = .03, p >.05.
Sheffield et al. (2021) USA	Secondary data analysis	Comparison group: none; Sample size: 814; Setting: Community Age: Range: 18-85; Gender: ND from larger study sample; Ethnicity: ND from larger study sample; Diagnosis: 5 (0.6%) psychotic disorder.	PDI-21	Negative self-beliefs: BDI-II (nine negative cognition items); ASR (six items from the Anxious/Depressed sub-scale consistent with negative self-beliefs) Worry: PSWQ	Negative self-beliefs were significantly associated with GD (n =439, r = 0.09, p = 0.07), indicating that worse negative self-beliefs were associated with worse delusional ideation. Worry was not significantly related to GD using the PSWQ (n = 228, r = .04, p = .53.

Smith et al. (2005) UK	CS	Comparison group: HC; Sample size: GD: 20; HC = 21 control; Setting: INP; Community (CMHT); GP; Mean age: GD = 37.1 (10.1); HC =33.1 (10.8); Gender: GD: 14 Male, 6 Female; HC: 12 Male, 9 Female; Ethnicity: NR; Diagnosis: GD: SZ (n=12); SAZ (n=4); BAD (n=4) - bipolar affective disorder. Co-occurring other delusion types.	SCAN (sections 17, 18, and 19) semi structured interview of positive symptoms of psychosis;	Depression: BDI; Anxiety: BAI; Self-esteem: RSCQ	No significant differences between GD vs control in RSCQ scores (t [39] = 0.41; p = .683), BDI scores (t [39]= -0.18; p = .862), or BAI scores (t [39] = 1.20; p = .236).
Smith et al. (2006) UK	CS	Comparison group: none; Sample size: 100 (17% GD); Setting: INP, OP; Mean age: 39 (10.9); Gender: M 68%; Ethnicity: Almost 70% described themselves as White-British, 10% as Black- Caribbean, 7% as Black-African and 11% as from other ethnic background Diagnosis: 78% SZ, 20% SAZ, 2% of DD; PRP trial.	SAPS	Depression: BDI Self-esteem: RSES	GD inversely associated with depression and low self-esteem (GD and BDI r_s = -0.38, p < .001; and RSES r_s = -0.33, p < .001) Ordered logistic estimations (logistic regression) for GD and: Depression OR= -.108 SE=.037 z = -2.89, p=.004; Self-esteem OR= -.015 SE= .060 z = -.25 p= .804.
Therman et al. (2014) Sweden	CS	Comparison group: none; Sample size: 31,822; Setting: Community GP; Mean age: mean age 51.4 (range	CAPE (GD: "being important, being special") items	Anxiety: Questions related to DSM-IV criteria for GAD	GD only slightly increased with GAD and MDE. No p values reported. The odds ratios for the top 5% high-scorers

		41-61); Gender: 100% F; Ethnicity: NR; Diagnosis: None.		Depression: Questions related to DSM-IV criteria for a MDE Lifetime diagnosis	of GD having either disorder was 2.0 (95% CI [1.8-2.2].
Ullrich et al. (2014) USA	Cohort	Comparison group: none; Sample size: 1136, 328 (28.9%) were deluded at baseline; Setting: INP Age: 29.7 (6.2) (range 18-40); Gender: Male (n = 667, 58.7%); Ethnicity: White (n = 785, 69.1%); Diagnosis: Nonaffective psychosis (n = 245, 21.6%), affective disorder including depression and bipolar disorder (n = 596, 52.5%), substance abuse/dependence (n = 274, 24.0%), and personality disorder (n = 21, 1.9%).	Interviewer rated based on questionnaire. GD question: Having special gifts/powers	Anger and elation both: MMADS	Anger significantly related to and having special gifts/powers (AOR = 2.98, 95% CI [1.67, 5.30]) $p < .05$. Elation significantly related to having special gifts/powers (AOR = 65.42, 95% CI = [35.05, 122.10], $p < .05$.
Van Putten et al. (1976) USA	CS	Comparison group: none; Sample size: Total= 59; GD= 13; Setting: INP, OP; Mean age: NR; Gender: NR; Ethnicity: NR; Diagnosis: SZ	GD was rated by one researcher as a part of BPRS	Depression, and anxiety both: researcher rated as a part of the BPRS	GD is significantly negatively correlated with anxiety ($r = -.59$, $p < .0005$) and depression ($r = -.57$), no p value reported.

Warman et al. (2010) USA	CS	Comparison group: none; Sample size: 121; Setting: Community GP (students); Age: 19.32 (1.74) (Range: 18 to 22); Gender: Female 73%; Ethnicity: White (85%); Diagnosis: Never received a diagnosis of SZ, SAZ, SZF or DD.	PDI-40	Self-esteem: RSES	No relationship emerged between self-esteem and GD r = -.06, p = .52.
Warman & Lysaker (2011) USA	CS	Comparison group: none; Sample size: 30; Setting: OP; Mean age: 48.93 (5.11) (Range 33 - 58); Gender: 100% M; Ethnicity: 50% White; Diagnosis: all SZ or SAZ.	PANSS, PDI-40	Self-esteem: MSEI	No relationships emerged between global self-esteem (MSEI total score) and: GD (PDI) (r = .05, p = .81) or PANSS GD (r= -.02, p = .89).

Note. ASR = Achenbach Adult Self-Report (Achenbach & Rescorla, 2003); BAI = Beck Anxiety Inventory (Beck et al., 1988); BDI-II = Beck Depression Inventory (Beck et al., 1961); BPRS = Brief Psychiatric Rating Scale (Overall & Gorham, 1962); CAPE = Community Assessment of Psychic Experiences (Stefanis et al., 2002); CDSS = Calgary Depression Scale (Addington et al., 1993); CDSS = Calgary Depression Scale for Schizophrenia (Addington et al., 1993); CES-D = Center for Epidemiologic Studies Depression Scale (Radloff, 1977); CMHT = Community Mental Health Team; DASS-21 = Depression, anxiety and stress scales (Lovibond & Lovibond, 1995); DD = delusion disorder; DOAI = Delusion and its Origin Assessment Interview (Hayashi et al., 2021); DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (APA, 2006); GAD = Generalized Anxiety Disorder; (GP = general population; HADS = Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983); HC = healthy controls; IAPT = improving access to psychological therapy; INP = inpatient; MADRS = Montgomery-Asberg Depression Rating Scale (Montgomery & Åsberg, 1979); MD = Mood disorder; MDE = Major Depressive Episode; MESI = Multidimensional Self-Esteem Inventory (O'Brien and Epstein, 1998); MIDAS = Motivational Interventions for Drug and Alcohol Misuse in Schizophrenia Trial; MMADS =MacArthur-Maudsley Assessment of Delusions Schedule (Appelbaum et al., 1999); OP = outpatient; PANSS = Positive and the Negative Syndrome Scale (Kay et al., 1987); PC = Paranoia Checklist (Freeman et al., 2005); PDI-21 = Peters Delusion Inventory (Peters et al., 2004); PDI-40 = Peters et al. Delusions Inventory (Peters et al., 1999); PRP = Prevention of Relapse in Psychosis PRP trial; PSWQ = Penn State Worry Questionnaire (Meyer et al., 1990); PSYRATS = Psychotic Symptom Rating Scales (Haddock et al., 1999); RSCQ = Robson Self Concept Questionnaire (RSCQ; Robson, 1989); RSES = Rosenberg Self-Esteem Scale (Rosenberg, 1965); SAPS= Scale for the Assessment of Positive Symptoms (Andreasen, 1984); SAZ schizoaffective disorder; SCAN = Schedules for Clinical Assessment in Neuropsychiatry (version 1.0; World Health Organization, 1994); SCID-I CV = Structured Clinical Interview for DSM-IV Axis I Disorder, clinical version (First et al., 1997); SERS-SF = Self-Esteem Rating Scale-Short Form (Lecomte et al., 2006); SZ= schizophrenia; SZF = schizophreniform disorder

Publication dates varied across included papers from 1976 (Van Putten et al., 1976) to 2021 (Sheffield et al., 2021). Regarding study design, one study involved secondary data analysis (Sheffield et al., 2021), two studies had a cohort design (Schennach et al., 2015; Ullrich et al., 2014), one had a mixed cohort and cross-sectional design (Moritz et al., 2010), and the remaining 17 studies adopted a cross-sectional design. Seven studies were conducted in America (Appelbaum et al., 1999; Ben-Zeev et al., 2012; Sheffield et al., 2021; Ullrich et al., 2014; Van Putten et al., 1976; Warman & Lysaker, 2011; Warman et al., 2010); seven in the United Kingdom (Boyden et al., 2015; Fowler et al., 2006; Garety et al., 2013; Hartley et al., 2012; Jacobsen et al., 2019; Smith et al., 2006; Smith et al., 2005), three in Germany (Moritz et al., 2010; Schennach et al., 2015; Schlier et al., 2015), and one each in France (Bortolon et al., 2019), Japan (Hayashi et al., 2021), Spain (de Portugal et al., 2009), and Sweden (Therman et al., 2014).

The total number of participants across all studies was 37,873 (mean = 1,803, median = 121), however, three studies used participants from the same trial (Fowler et al., 2006; Garety et al., 2013; Smith et al., 2006). The decision was made to include papers with potentially overlapping data sets given the studies were carried out on sub-samples of a larger trial sample and had differing hypotheses. Total sample size varied from 30 (Warman & Lysaker, 2011) to 31,822 (Therman et al., 2014), however, the sample size for a specific GD sub-sample was as small as n=4 (de Portugal et al., 2009).

The majority of the studies involved a patient population in which GD was measured however three were general population samples (Sheffield et al., 2021; Therman et al., 2014; Warman et al., 2010), specifying either no diagnosis or only 0.6% having a diagnosis of psychotic disorder (Sheffield et al., 2021). The remaining studies had diagnoses typically associated with delusions with the majority including schizophrenia or schizoaffective

disorder within samples (n=15), excluding one study which did not report on diagnoses (Appelbaum et al., 1999).

Twelve studies did not offer a comparison group to GD, one study compared with depressed controls (Boyden et al., 2015), one study compared GD with other Delusion Disorder (DD) types (de Portugal et al., 2009), two studies compared GD with persecutory delusions (Garety et al., 2013; Jacobsen et al., 2019), and five studies compared with a "healthy control" equivalent (Appelbaum et al., 1999; Fowler et al., 2006; Moritz et al., 2010; Schlier et al., 2015; Smith et al., 2005).

Eleven studies did not report on the ethnicity of included participants, for those that did, nine reported a majority White sample (Ben-Zeev et al., 2012; Boyden et al., 2015; Fowler et al., 2006; Garety et al., 2013; Hartley et al., 2012; Smith et al., 2006; Ullrich et al., 2014; Warman & Lysaker, 2011; Warman et al., 2010). Reported mean age of samples ranged from 19.32 (Warman et al., 2010) to 54 (de Portugal et al., 2009) years old, with the widest age range of 18-85 years old (Sheffield et al., 2021). Three studies did not provide sufficient detail to determine gender split (Appelbaum et al., 1999; Sheffield et al., 2021; Van Putten et al., 1976), for the remaining studies the mean % of males was 61.2% and fluctuated from 0% male (Therman et al., 2014) to 100% male (Warman & Lysaker, 2011).

Measures of Grandiose Delusions

There was large variability in the approach and measures used to determine the presence of GD, with the majority of studies opting for a minimum of two types of measures/approaches, excluding five studies which used a single measure only (Ben-Zeev et al., 2012; Bortolon et al., 2019; Moritz et al., 2010; Schennach et al., 2015; Smith et al., 2005; Therman et al., 2014; Warman et al., 2010). One study specified assignment to DD type being clinician-rated based on the description provided by participants without an

accompanying measure mentioned (Jacobsen et al., 2019). All included studies relied on self-reports from participants.

Some measures were used across multiple studies including the Positive and the Negative Syndrome Scale (PANSS; Kay et al., 1987), which was used in eight studies (Bortolon et al., 2019; Fowler et al., 2006; Hartley et al., 2012; Hayashi et al., 2021; Moritz et al., 2010; Schennach et al., 2015; Schlier et al., 2015; Warman & Lysaker, 2011) and has a moderate reliability for the positive scale: ICC=.62; Peralta & Cuesta, 1994); the Psychotic Symptom Rating Scale (PSYRATS; good reliability ICC=.88; Haddock et al., 1999) was used in two studies (Ben-Zeev et al., 2012; Hartley et al., 2012); the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984) in two studies (Garety et al., 2013; Smith et al., 2006); the Peters Delusion Inventory- 21 item (PDI-21; Peters et al., 2004) and the 40 items (PDI-40; Peters et al., 1999) were used in two (Boyden et al., 2015; Sheffield et al., 2021) and three studies (Fowler et al., 2006; Warman & Lysaker, 2011; Warman et al., 2010), respectively. Good internal consistencies are found for both the PDI-21(α=.82; Peters et al., 2004) and the PDI-40 (α=.88; Peters et al., 1999).

Other measures used in single studies only were the MacArthur-Maudsley Assessment of Delusions Schedule interview (MMADS; Appelbaum et al., 1999); the Psychosis module of the Structured Clinical Interview for DSM-IV Axis I Disorder, Clinical Version (SCID-I CV; First et al., 1997); the Delusion and its Origin Assessment Interview (DOAI; Hayashi et al., 2021); the Paranoia Checklist (PC; excellent internal consistency α>0.9; Freeman et al., 2005) used in Schlier et al. (2015); the Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962) used in Van Putten et al. (1976), with a good internal consistency (α=.87); the Schedules for Clinical Assessment in Neuropsychiatry (SCAN; version 1.0; World Health Organization, 1994) used in Smith et al. (2005); and the Community Assessment of Psychic

Experiences (CAPE; Stefanis et al., 2002) used in Therman et al., (2014) which has good internal consistency (α=.84; Mark & Toulopoulou, 2016).

Emotion Measures

A range of emotion factors were identified in relation to GD such as depression, anxiety/stress/worry, self-esteem, elation, and anger. Some emotion factors encompassed multiple sub-emotions and are presented separately (negative affect, depressed factor and negative self-belief). Eleven studies reported on more than one emotion factor and corresponding information is presented separately for each emotion factor.

Depression

Measure of depression varied across studies and one study measured sadness which has been grouped under the wider depression label. All studies used a form of self-report as the main measure of depression. All studies used a single measure of depression excluding one where two measures were used to measure depression (Schlier et al., 2015).

Across the studies, eight different validated measures were used. The Beck Depression Inventory (BDI-II; Beck et al., 1961) was used in four studies (Garety et al., 2013; Schlier et al., 2015; Smith et al., 2006; Smith et al., 2005) and has excellent internal consistency (α=0.9; Wang & Gorenstein, 2013); the Calgary Depression Scale for Schizophrenia (CDSS; Addington et al., 1993) was used in two studies (Bortolon et al., 2019; Schennach et al., 2015) and has acceptable internal consistency (α=.76; Müller et al, 2005); the Depression, Anxiety and Stress Scales (DASS-21; excellent internal consistency depression scale α=.91; Lovibond & Lovibond, 1995) was used by one study (Jacobsen et al., 2019); the Hospital Anxiety and Depression Scale (HADS; good internal consistency α=.83; Zigmond & Snaith, 1983) was used by one study (Boyden et al., 2015); the Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979) was used by one study (de Portugal et al., 2009) and has good internal consistency (α=.84; Fantino & Moore,

2009); the Center for Epidemiologic Studies Depression Scale (CES-D; good internal consistency α=.85; Radloff, 1977) was used by one study (Schlier et al., 2015); the Brief Psychiatric Rating Scale (BPRS) was used by one study (Van Putten et al., 1976); and the PANSS was used in one study (Hartley et al., 2012) which has a good reliability for the negative scale: ICC=.8; Peralta & Cuesta, 1994).

Two studies did not use a validated measure. One study used a set of questions related to DSM-IV criteria for major depressive episode (Therman et al., 2014) and one study used a single item "*How sad do you feel right now*?" (Ben-Zeev et al., 2012).

Anxiety

Measure of anxiety was varied across studies with two studies exploring stress and worry which have been grouped under the anxiety label for the purpose of the review. All studies used a form of self-report as the main measure of anxiety, with all using a single measure.

Across the studies, a mix of five different validated measures were used. The Beck Anxiety Inventory (BAI; Beck et al., 1988) was used by two studies (Garety et al., 2013; Smith et al., 2005) and has excellent internal consistency (α=.94; Fydrich, Dowdall & Chambless, 1992); the HADS was used by one study (Boyden et al., 2015); the PANSS was used by one study (Hartley et al., 2012); the BPRS was used by one study (Van Putten et al., 1976); and the DASS-21 was used by one study (Jacobsen et al., 2019)

Two studies did not use a validated measure. One study used a set of questions related to DSM-IV criteria for generalised anxiety disorder (Therman et al., 2014) and one study used a single item "*How anxious do you feel right now*?" (Ben-Zeev et al., 2012).

Stress. One study reported on the factor of stress separately in addition to anxiety using the DASS-21 (Jacobsen et al., 2019).

Worry. The Penn State Worry Questionnaire (PSWQ; excellent internal consistency α=.95; Meyer et al., 1990), a worry-specific measure was used by one study (Sheffield et al., 2021).

Self-Esteem

Measure of self-esteem varied across studies. All studies used a single form of self-report as the main measure of self-esteem.

Across the studies, a mix of four different validated measures were used. The Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) was the most frequently used, adopted in five studies (Fowler et al., 2006; Garety et al., 2013; Moritz et al., 2010; Smith et al., 2006; Warman et al., 2010) and has a good internal consistency (α=.84-.86; Tinakon & Nahathai, 2012). Other measures used were the Self-Esteem Rating Scale-Short Form (SERS-SF; good internal consistency α=.87; Lecomte et al., 2006) negative sub-scale score (Ben-Zeev et al., 2012); the Robson Self Concept Questionnaire (RSCQ; Robson, 1989; used by Smith et al., 2005) which has excellent internal consistency (α=.94-.97; Ghaderi, 2005); and the Multidimensional Self-Esteem Inventory (MSEI; O'Brien & Epstein, 1988; used by Warman & Lysaker, 2011).

One study reported on negative self-belief using a measure comprised of the nine negative cognition items on the BDI and six items from the anxious/depressed sub-scale of the Achenbach Adult Self-Report (ASR; Achenbach & Rescorla, 2003) which were consistent with negative self-beliefs (Sheffield et al., 2021).

Elation

Elation was explored using the MMADS in one study (Ullrich et al., 2014).

Anger

Anger was explored using the MMADS in one study (Ullrich et al., 2014).

Mixed Emotion Factors

Negative Affect. Two studies used a negative affect name for emotion factors with a differing mix of sub-emotions. One study used negative affect as an umbrella term for unhappy, frightened, anxious, or angry using the MMADS as a measure (Appelbaum et al., 1999). One study used negative affect as an umbrella term for anxiety, tension, unpleasantness, excitement and anger, using the DOAI as a measure (Hayashi et al., 2021).

Depressed Factor. One study reported on a 'depressed factor' comprising anxiety, guilt feelings and depression using the PANSS as a measure (Hayashi et al., 2021).

Study Quality

A summary of quality assessment ratings for all 21 included studies is provided in Table 3. No papers were excluded from the narrative synthesis based on poor quality.

Across the 14 criteria items, one of the included studies adequately met four items (Appelbaum et al., 1999); two studies adequately met five items (Bortolon et al., 2019; Therman et al., 2014); five studies adequately met six items (Hartley et al., 2012; Schlier et al., 2015; Sheffield et al., 2021; Warman & Lysaker, 2011; Warman et al., 2010); four studies adequately met seven items (Boyden et al., 2015; de Portugal et al., 2009; Hayashi et al., 2021; Van Putten et al., 1976); two studies adequately met eight items (Moritz et al., 2010; Smith et al., 2006); and the remaining seven studies adequately met nine items.

Items 1 and 4 received the highest number of items rated as adequately addressed. Item 1 related to having a clear research question/aim was rated as adequately addressed for 19 studies with only two studies receiving a partially addressed rating (Appelbaum et al., 1999; de Portugal et al., 2009). Similarly, item 4, related to the recruitment of the sample population was rated as adequately addressed for 19 studies, with one study receiving a

partial rating (Schlier et al., 2015), and one study where this was not reported (Boyden et al., 2015).

No studies adequately addressed the item about sufficient detail around power calculations concerning sample sizes and whether sufficient power was achieved to identify significant results, however, four papers received a partially adequate rating (Ben-Zeev et al., 2012; Garety et al., 2013; Hartley et al., 2012; Sheffield et al., 2021). Only one study received an adequately addressed rating for item 12 relating to blinding to the exposure status of assessors (Van Putten et al., 1976), with most studies not reporting this. Similarly, only one study received an adequately addressed rating for item 13 regarding loss to follow-up from baseline (Fowler et al., 2006), three studies received a not adequately addressed rating (Moritz et al., 2010; Schennach et al., 2015; Ullrich et al., 2014), but the majority received an NA rating due to study design.

Most studies received a rating of adequately addressing item 9 related to defining and using a valid and reliable measure of GD with only four studies receiving a partial rating (Appelbaum et al., 1999; Hartley et al., 2012; Ullrich et al., 2014; Van Putten et al., 1976). This was similarly reflected in item 11 related to defining and using a valid and reliable measure of an emotion factor(s). Fifteen studies received an adequately addressed rating excluding six studies which received a partially addressed rating (Appelbaum et al., 1999; de Portugal et al., 2009; Hartley et al., 2012; Hayashi et al., 2021; Therman et al., 2014; Van Putten et al., 1976).

However, when considering ratings for item 14, relating to statistically adjusting for key confounding variables, methodological quality was less strong. Only six studies adequately addressed this item (Garety et al., 2013; Hartley et al., 2012; Sheffield et al., 2021; Smith et al., 2006; Smith et al., 2005; Ullrich et al., 2014), with a further four being

rated as inadequately addressing (de Portugal et al., 2009; Therman et al., 2014; Warman & Lysaker, 2011; Warman et al., 2010), and the remaining 11 receiving a partial rating.

Table 3

Methodological Quality Assessment Ratings for Included Studies (Sorted Alphabetically by First Author)

Author and year	1. Clear research question or objective	2. Clearly specified and defined study population	3. Participation rate of eligible >50%	4. Recruited from same/similar populations? Uniformly applied inclusion and exclusion criteria	5. Sample size justification/ powered analyses	6. GDs measured prior to the emotional factor(s)	7.Sufficient timeframe to assess association between GDs and emotion factor(s)	8. Examination of variations/degree in GDs related to the emotion measure(s)	9. Clearly defined, valid, reliable, consistently applied GD measures	10.GD assessed more than once over time	11. Clearly defined, valid, reliable, consistently applied emotion measure(s)	12. Outcome assessors blinded to the GD status of participants	13.Loss to follow-up after baseline <20%	14. Key potential confounding variables measured and adjusted statistically for
Appelbaum et al. (1999)	P	P	Yes	Yes	No	Yes	No	Yes	P	No	P	NR	NA	P
Ben-Zeev et al. (2012)	Yes	Yes	Yes	Yes	P	Yes	NR	Yes	Yes	Yes	Yes	No	NA	P
Bortolon et al. (2019)	Yes	Yes	NR	Yes	No	No	No	P	Yes	No	Yes	No	NA	P
Boyden et al. (2015)	Yes	Yes	NR	NR	No	Yes	Yes	P	Yes	Yes	Yes	NR	NA	P
de Portugal et al. (2009)	P	Yes	Yes	Yes	No	Yes	Yes	P	Yes	Yes	P	No	NA	No
Fowler et al. (2006)	Yes	Yes	Yes	Yes	No	Yes	Yes	P	Yes	P	Yes	NR	Yes	P
Garety et al. (2013)	Yes	Yes	Yes	Yes	P	Yes	Yes	P	Yes	No	Yes	NR	NA	Yes

Hartley et al. (2012)	Yes	Yes	Yes	Yes	P	No	No	Yes	P	No	P	No	NA	Yes
Hayashi et al. (2021)	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	P	No	NA	P
Jacobsen et al. (2019)	Yes	Yes	Yes	Yes	No	Yes	Yes	P	Yes	Yes	Yes	NR	NA	P
Moritz et al. (2010)	Yes	Yes	NR	Yes	No	Yes	Yes	P	Yes	Yes	Yes	NR	No	P
Schennach et al. (2015)	Yes	Yes	Yes	Yes	No	Yes	Yes	P	Yes	Yes	Yes	NR	No	P
Schlier et al. (2015)	Yes	P	Yes	P	No	Yes	Yes	P	Yes	No	Yes	No	NA	P
Sheffield et al. (2021)	Yes	P	NR	Yes	P	NR	Yes	P	Yes	NR	Yes	NR	NA	Yes
Smith et al. (2005)	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	NR	NA	Yes
Smith et al. (2006)	Yes	Yes	Yes	Yes	No	NR	Yes	P	Yes	No	Yes	No	NA	Yes
Therman et al. (2014)	Yes	Yes	Yes	Yes	No	No	No	P	Yes	No	P	No	NA	No
Ullrich et al. (2014)	Yes	Yes	Yes	Yes	No	Yes	Yes	P	P	Yes	Yes	No	No	Yes
Van Putten et al. (1976)	Yes	No	Yes	Yes	No	Yes	Yes	P	P	Yes	P	Yes	NA	P
Warman et al. (2010)	Yes	P	NR	Yes	No	Yes	Yes	P	Yes	No	Yes	NR	NA	No
Warman & Lysaker (2011)	Yes	P	Yes	Yes	No	Yes	NR	P	Yes	No	Yes	NR	NA	No

Note. GD = grandiose delusions, P = partially achieved, NR = not reported, NA = not applicable

Main Findings

Depression

Thirteen studies explored depression and sadness in relation to GD, all using a cross-sectional design, excluding one which used a naturalistic/cohort design (Schennach et al., 2015). Nine studies found GD to be significantly related to lower depression either with the presence or absence of GD or when compared to a control group, all using patient samples. Five studies had no comparison group (Ben-Zeev et al., 2012; Hartley et al., 2012; Schennach et al., 2015; Smith et al., 2006; Van Putten et al., 1976) and reported relationship strength ranged from r= -.14 (Ben-Zeev et al., 2012) to r= -.57 (Van Putten et al., 1976); one study found depression to be lower for GD compared with depressed controls, p<.0001 (Boyden et al., 2015); one found depression to be lower in GD vs other DD types, p=.048 (de Portugal et al., 2009); and two studies found depression to be lower in GD compared to persecutory delusions, p<.001 and p=.05 (Garety et al., 2013; Jacobsen et al., 2019).

Of the four studies with non-significant findings, all excluding one (Therman et al., 2014) were in patient samples, with two looking at differences for patients with GD compared to healthy controls (Schlier et al., 2015; Smith et al., 2005) and one looking only at correlations (Bortolon et al., 2019).

There was mixed evidence of the relationship of depression with GD and no studies allowed for causality to be established. The direction of relationships remained negative in all significant findings however strength varied with the strongest correlations being found by Van Putten et al. (1976), however, this was in a smaller sample size. Overall, there appeared to be more support in favour of lower levels of depression associated with GD but contradictory evidence was present particularly when there were healthy controls.

Anxiety

Eight studies reported on the relationship between GD and anxiety all using cross-sectional designs.

Four studies found a significant negative relationship between GD and anxiety, all within patient samples. Van Putten et al. (1976) found a correlation of $r= -.59$, $p<.0005$, and Hartley et al. (2012) found anxiety to be significantly lower when GD were present compared to absent ($p=.045$). When comparing to controls Boyden et al. (2015) found lower anxiety in GD compared to depressed controls ($p<.0001$); Garety et al. (2013) found lower anxiety in GD compared to persecutory delusions ($p<.001$);

Contrastingly, four studies did not find a significant relationship. Therman et al. (2014) found GD only slightly increased with the presence of GAD in non-patient samples, and Ben-Zeev et al. (2012) found anxiety did not significantly predict GD occurrence in patient samples. Similarly, no significant differences were found when compared to persecutory delusion (Jacobsen et al., 2019) or healthy controls (Smith et al., 2005).

Reporting specifically on stress within patient samples (total n=41) Jacobsen et al. (2019) found no significant difference in stress levels associated with GD compared to persecutory delusions. For worry specifically, Sheffield et al. (2021) found no significant relationship with GD within non-patient community samples (GD n=439).

Overall, findings are not clear. There is some evidence to support a possible link between lower anxiety associated with GD however this has not been found consistently even with the uniform use of patient samples across studies. Additionally, no studies found a relationship between GD and increased anxiety. All studies used a cross-sectional design and as previously, there is a dearth of research allowing for longitudinal conclusions to be drawn.

Self-Esteem

Eight studies explored the relationship between GD and self-esteem all using cross-sectional designs, excluding Moritz et al. (2010) who instead utilised a cohort design.

Two studies found a significant negative relationship between GD and self-esteem. Within a patient sample, Ben-Zeev et al. (2012) found lower self-esteem significantly predicted the occurrence of GD (r= .04 (SE=0.02), OR=1.04, 95% CI [1.01, 1.08], p<.05). Similarly, but reporting for their non-clinical student sample, Fowler et al. (2006) found GD had a small weak association with lower self-esteem (bivariate r= -.2; p=.014).

Two studies found contrasting results. Garety et al. (2013) found self-esteem to be higher in patients with GD however this was compared to patients with persecutory delusions (OR=0.82 OR=0.82, p<.0001). Moritz et al. (2010) also explored GD in a patient sample and found GD to be significantly but modestly positively correlated with self-esteem longitudinally (r=.26, p<.05).

Smith et al. (2006) within a small patient sample (GD n=17) initially found a significant negative relationship between GD and low self-esteem (rs= -.33, p<.001), like Ben-Zeev et al. (2012) and Fowler et al. (2006). However, following ordered logistic regression this effect disappeared (OR= -.015, SE=.060, z= -.25 p=.804). Three studies did not find any significant relationship, all being cross-sectional designs, two looking at GD in patient samples (Smith et al., 2005; Warman & Lysaker, 2011), and one in student samples (Warman et al., 2010). Smith et al. (2005) compared individuals with GD with healthy controls.

Overall, the findings regarding self-esteem and GD were inconclusive and somewhat contradictory. Most studies used a cross-sectional design and even for significant findings

directions of associations were split with patient samples finding both positive and negative relationships between self-esteem and GD.

Elation and Anger

Ullrich et al. (2014) found both elation and anger were significantly associated with GD in patient samples using a cohort design (delusions present n=1136). No other studies explored these factors to allow for comparisons of findings.

Mixed Emotion Factors

No studies measured mixed factors in the same way as any other so findings are presented separately. All studies used a cross-sectional design excluding Sheffield et al. (2021) conducting secondary data analyses.

Hayashi et al. (2021) found a significant negative relationship between GD and negative affect ($r = -.421$, $p<.01$) and "depressed factor" (anxiety, guilt feelings depression; $r = -.209$, $p<.05$) within patient samples with psychosis diagnoses (total sample: n=108). Similarly, Appelbaum et al. (1999) found negative affect to be significantly lower in individuals with GD compared to those without ($p<.001$) within patient samples (GD n=141). These align with the findings above of studies demonstrating support for lower levels of depression being associated with GD.

Sheffield et al. (2021) found "negative self-belief" to be significantly related to GD ($r=.09$, $p=.07$) within non-patient community samples (GD n=439). These may parallel studies finding support for lower self-esteem being associated with GD.

Discussion

The present review aimed to better understand the relationship between GD and emotion factors in the hopes of casting light on potential mechanisms involved in the formation and maintenance GD. The review found some support for the link of GD with

emotion factors; however, overall findings were mixed. Despite larger numbers of papers identified within initial searches, many were excluded following screening due to not focusing on GD specifically in relation to emotion factors. This is in line with previous literature commenting on a dearth of evidence in this area (Grbic, 2013). In total eight emotion factors were identified from studies with some being composites of multiple sub-emotion factors. The majority of studies explored the relationship between three main emotion factors: depression, anxiety and self-esteem.

Findings of the relationship of GD with lower levels of depression were perhaps the most convincing with a majority of significant findings in patient samples. However, some studies did find conflicting evidence particularly when GD were compared to healthy controls adopting a good sample size (Schlier et al., 2015). However, it is important to note two studies did not provide exact p values for their non-significant results. These findings seem to be in opposition to those in the literature commenting on the often co-occurrence of GD with persecutory delusions and depression (Knowles et al., 2011).

Findings for anxiety provided even less confidence in potential links with GD. Some studies found lower anxiety associated with GD, however, there appeared to be a greater number of studies finding no significant relationship, particularly when worry and stress were included under the anxiety umbrella and no studies reported a link of GD with increased anxiety. Additionally, for studies reporting significant findings, study methodology casts further doubt. Boyden et al. (2015) compared GD to depressed controls. Lower anxiety in the GD group may have resulted from the comorbidity of anxiety and depression that often occurs (Saha et al., 2021; Wilson et al., 2020). Garety et al. (2013) compared anxiety in GD vs persecutory delusions, where again literature has suggested increased anxiety to be related to persecutory delusions (Bennetts et al., 2022). Other significant findings were within small sample sizes (GD n=13; Van Putten et al., 1976) or in papers with poorer methodological

quality (Hartley et al., 2012). Overall, a relationship between GD and anxiety cannot be concluded based on the current review. These findings are in contrast to the relationship found between higher levels of anxiety and other delusion types such as persecutory delusions (Bennetts et al., 2022).

Findings for self-esteem were more convoluted with conclusions difficult to ascertain. Both significant and non-significant findings were present, and even within significant findings, the direction of the relationships was contradictory, with findings for both increased and decreased self-esteem. Overall, study quality was slightly better for studies that found a significant negative relationship between self-esteem and GD as opposed to a positive relationship, however, one was in a student population. More convincingly Moritz et al. (2010) looked longitudinally in patient samples and found a positive relationship. Overall, it is not possible to state the relationship between self-esteem and GD based on the current review findings. The inconsistency in findings matches those found in the previous review, Grbic (2013) suggested a possible role of mood on the relationship, which we were not able to assess within the present review. Additionally, Knowles et al. (2011) comment on the conflicting evidence within the literature and potential support for both DAD and emotion-consistent theories of GD. They also suggested a role of unstable self-esteem, with both higher and lower self-esteem impacting GD within their proposed model. It may be that self-esteem fluctuates for individuals experiencing GD, an occurrence which has been found to happen within bipolar affective disorder, where GD are also commonly reported (Bentall et al., 2011; Knowles et al., 2007).

Concerning DAD and emotion-consistent theories, our findings of lower levels of depression associated with GD and no studies linking higher anxiety with GD, appear to be in contrast to DAD theories, where we might expect GD to be present as a protective element for the presence of negative feelings such as depression and anxiety. Further support for a

contrast of our findings with DAD theories, GD were found to be associated with less negative affect, increased elation, and found to have an inverse relationship with "depressed factor" and anger all in patient samples. In contrast, negative self-belief was significantly related to GD, however, this was in non-patient samples which may explain the conflicting findings. Overall, this could suggest more support for emotion-consistent theories of GD but due to limitations of the current evidence base, it is not possible to say for certain.

Strengths and Limitations: Current Review

The review was completed in adherence with PRISMA guidelines and pre-registered increasing scientific rigour. Three separate databases were searched aligning with recommendations (Siddaway et al., 2019) and forward and backward citation searching were also conducted increasing confidence that all relevant papers were identified. The review also employed a second reviewer at each stage of the study selection process and 100% of included studies were independently assessed for methodological quality and data extraction by a second reviewer, increasing interrater reliability and reducing risk of bias. A formal singular appraisal tool was used to assess the quality of all papers allowing for better comparisons of methodological quality across included studies and increasing the reliability of ratings.

However, there were limitations to the present review. Due to large heterogeneity between measures within studies and overlapping samples, a meta-analysis was not possible. The validity of findings from narrative syntheses alone has been questioned given a suggestion this approach is more vulnerable to subjectivity bias (Campbell et al., 2018). The review focussed only on studies with distinguishable adult samples which excluded some papers. Publication bias may also be present due to the decision not to search or include studies from grey literature. Additionally, included studies were limited to those available in English, introducing a potential language bias. These potentially result in relevant studies

being missed, impacting possible conclusions and generalisability. The current review included quantitative studies only reducing the number of eligible studies. Future reviews should explore the current qualitative evidence base concerning the factors associated with GD.

Research Limitations, and Recommendations

The majority of studies explored the relationship between GD and emotions through correlations in cross-sectional designs, meaning causality was not possible to establish. Additionally, a range of approaches were used to measure both GD and emotion factors including standardised and unstandardised measures. This heterogeneity meant that meaningful comparisons and conclusions were difficult to establish. Future research should focus on study designs that would allow causality to be explored such as longitudinal and experimental studies. Involving at-risk and patient populations within study samples would further support the exploration of factors associated with the formation of GD.

Of the included studies, many involved patient samples leading to increased confidence in conclusions. Some papers compared with controls to further determine differences in relationships for GD, however, these were often student samples for “healthy controls” which may not be representative of the wider population and limits applicability. Future research should explore differences in relationships with matched controls to minimise potential confounding factors between groups as often confounding variables were not considered within analyses. Additionally given patient samples within studies were sometimes mixed, exploring emotion factors related to GD in different patient groups (e.g., schizophrenia vs bipolar affective disorder) could be beneficial.

The methodological quality of studies varied and although total sample sizes were sometimes larger, GD sample sizes specifically were either not provided or often a much smaller proportion of the total sample reducing confidence in the representativeness of

findings to wider populations. Very few studies set out to specifically explore GD as a focus and some studies neglected to control for the co-occurrence of other delusion types which may have confounded findings. GD have been suggested to often occur alongside other delusion types such as persecutory delusions which may have distinct processes in their formation and maintenance (Garety et al., 2013; Knowles et al., 2011). Future research should include larger samples exploring the relationship between emotion factors and GD specifically.

Clinical Implications

It was hoped the findings would be beneficial in supporting the development of interventions specific to GD. However, given the inconsistency in findings, clinical implications are tentatively drawn. Given the differences between findings regarding GD and those previously found for other delusion types (persecutory), clinicians should specifically formulate according to the delusion subtype present. Additionally, collecting information on co-occurring symptoms and self-esteem within clinical practice would provide further insight into an individual's experience and the potential impact of GD. Findings also emphasise the importance of therapists keeping in mind potential barriers to engagement. The association of GD with more positive affect or the absence of negative affect may mean those seen within clinical practice are less driven towards change, given the positive impact of negative affect on help-seeking behaviour (Nagai, 2015).

Conclusion

Overall, better-quality research adopting different study designs is needed to better understand the role of emotion in the formation and maintenance of GD. Although there was support for the role of emotions particularly lower depression being linked with GD, the current evidence base was inconsistent on the whole. Given the discussed limitations caution should be applied within interpretation. It appears findings in this review regarding current

empirical evidence are similar to those in previous reviews, suggesting that this continues to remain methodologically limited despite the addition of new research. There remains insufficient empirical evidence to adequately understand the role of emotions in the formation and maintenance of GD, and this should remain a focus for future research.

References

Achenbach, T. M., & Rescorla, L. (2003). Manual for the ASEBA adult forms & profiles. In: University of Vermont, Research Center for Children, Youth, & Families

Addington, D., Addington, J., & Maticka-tyndale, E. (1993). Assessing depression in schizophrenia: The calgary depression scale. *British Journal of Psychiatry*, *163*(S22), 39-44. https://doi.org/10.1192/S0007125000292581

American Psychiatric Association. (2006). *Diagnostic and statistical manual of mental disorders* (4th ed.). https://doi.org/10.1176/ajp.152.8.1228

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Andreasen, N. C. (1984). Scale for the assessment of positive symptoms. *Psychiatrie & Psychobiologie*. https://doi.org/10.1037/t48377-000

Appelbaum, P. S., Robbins, P. C., & Roth, L. H. (1999). Dimensional approach to delusions: Comparison across types and diagnoses. *The American Journal of Psychiatry*, *156*(12), 1938-1943. https://doi.org/10.1176/ajp.156.12.1938

Bailey, T., Alvarez-Jimenez, M., Garcia-Sanchez, A. M., Hulbert, C., Barlow, E., & Bendall, S. (2018). Childhood trauma is associated with severity of hallucinations and delusions in psychotic disorders: A systematic review and meta-analysis. *Schizophrenia Bulletin*, *44*(5), 1111-1122. https://doi.org/10.1093/schbul/sbx161

Beck, A. T., & Rector, N. A. (2005). Cognitive approaches to schizophrenia: Theory and therapy. *Annual Review of Clinical Psychology*, *1*(1), 577-606. https://doi.org/10.1146/annurev.clinpsy.1.102803.144205

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893-897. https://doi.org/10.1037/0022-006X.56.6.893

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561-571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Ben-Zeev, D., Morris, S., Swendsen, J., & Granholm, E. (2012). Predicting the occurrence, conviction, distress, and disruption of different delusional experiences in the daily life of people with schizophrenia. *Schizophrenia Bulletin*, *38*(4), 826-837. https://doi.org/10.1093/schbul/sbq167

Bennetts, A., Stopa, L., & Newman-Taylor, K. (2022). What can experimental studies tell us about paranoia and anxiety? A systematic review with implications for theory and clinical practice. *Psychosis*, *14*(2), 162-175. https://doi.org/10.1080/17522439.2021.1909112

Bentall, R. P., Myin-Germeys, I., Smith, A., Knowles, R., Jones, S. H., Smith, T., & Tai, S. J. (2011). Hypomanic personality, stability of self-esteem and response styles to negative mood. *Clinical Psychology and Psychotherapy*, *18*(5), 397-410. https://doi.org/10.1002/cpp.780

Benzies, K. M., Premji, S., Hayden, K. A., & Serrett, K. (2006). State-of-the-evidence reviews: Advantages and challenges of including grey literature. *Worldviews on Evidence-Based Nursing*, *3*(2), 55-61. https://doi.org/10.1111/j.1741-6787.2006.00051.x

Boland, A., Cherry, M. G., & Dickson, R. (2017). *Doing a systematic review: A student's guide* (2nd ed.). SAGE.

Bortolon, C., Yazbek, H., Norton, J., Capdevielle, D., & Raffard, S. (2019). The contribution of optimism and hallucinations to grandiose delusions in individuals with schizophrenia. *Schizophrenia Research*, *210*, 203-206. https://doi.org/10.1016/j.schres.2018.12.037

Boyden, P., Knowles, R., Corcoran, R., Hamilton, S., & Rowse, G. (2015). A preliminary investigation into theory of mind and attributional style in adults with grandiose delusions. *Cognitive Neuropsychiatry*, *20*(2), 109-121. https://doi.org/10.1080/13546805.2014.974802

Campbell, M., Katikireddi, S. V., Sowden, A., McKenzie, J. E., & Thomson, H. (2018). Improving conduct and reporting of narrative synthesis of quantitative data (ICONS-Quant): Protocol for a mixed methods study to develop a reporting guideline. *BMJ open*, *8*(2), e020064. https://doi.org/10.1136/bmjopen-2017-020064

Collin, S., Rowse, G., Martinez, A. P., & Bentall, R. P. (2023). Delusions and the dilemmas of life: A systematic review and meta-analyses of the global literature on the prevalence of delusional themes in clinical groups. *Clinical Psychology Review*, *104*. https://doi.org/10.1016/j.cpr.2023.102303

Connors, M. H., & Halligan, P. W. (2020). Delusions and theories of belief. *Consciousness and Cognition*, *81*, 102935. https://doi.org/10.1016/j.concog.2020.102935

Corsini, R. (2001). *The dictionary of psychology* (1st ed.). Routledge. https://doi.org/10.4324/9781315781501

de Portugal, E., Gonzalez, N., Vilaplana, M., Haro, J. M., Usall, J., & Cervilla, J. A. (2009). An empirical study of psychosocial and clinical correlates of delusional disorder: the DELIREMP study. *Revista de Psiquiatria y Salud Mental*, *2*(2), 72-82. https://doi.org/10.1016/S1888-9891(09)72248-7

Fantino, B., & Moore, N. (2009). The self-reported Montgomery-Åsberg depression rating scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry*, *9*, 1-6. https://doi.org/10.1186/1471-244X-9-26

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1997). *SCID-I: User's guide for the structured clinical interview for DSM-IV axis I disorders*. American Psychiatric Press.

Fowler, D., Freeman, D., Smith, B., Kuipers, E., Bebbington, P., Bashforth, H., Coker, S., Hodgekins, J., Gracie, A., Dunn, G., & Garety, P. (2006). The brief core schema scales (BCSS): Psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples. *Psychological Medicine*, *36*(6), 749-759. https://doi.org/10.1017/S0033291706007355

Freeman, D., & Garety, P. A. (2003). Connecting neurosis and psychosis: The direct influence of emotion on delusions and hallucinations. *Behaviour Research and Therapy*, *41*(8), 923-947. https://doi.org/10.1016/S0005-7967(02)00104-3

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *British Journal of Psychiatry*, *186*(5), 427-435. https://doi.org/10.1192/bjp.186.5.427

Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders*, *6*(1), 55-61. https://doi.org/10.1016/0887-6185(92)90026-4

Garety, P. A., Gittins, M., Jolley, S., Bebbington, P., Dunn, G., Kuipers, E., Fowler, D., & Freeman, D. (2013). Differences in cognitive and emotional processes between persecutory and grandiose delusions. *Schizophrenia Bulletin*, *39*(3), 629-639. https://doi.org/10.1093/schbul/sbs059

Ghaderi, A. (2005). Psychometric properties of the self-concept questionnaire. *European Journal of Psychological Assessment*, *21*(2), 139-146. https://doi.org/10.1027/1015-5759.21.2.139

Gold, C., Erkkilä, J., & Crawford, M. J. (2012). Shifting effects in randomised controlled trials of complex interventions: A new kind of performance bias? *Acta psychiatrica Scandinavica*, *126*(5), 307-314. https://doi.org/10.1111/j.1600-0447.2012.01922.x

Grbic, S. (2013). *A systematic literature review of the role of self-esteem in persecutory and grandiose delusions and a grounded theory exploration of grandiose beliefs* [DClinPsy thesis, University of Sheffield]. White Rose eTheses Online. https://etheses.whiterose.ac.uk/4719/1/ethesis.pdf

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879-889. https://doi.org/10.1017/S0033291799008661

Harder, S. (2006). Self-image and outcome in first-episode psychosis. *Clinical Psychology & Psychotherapy*, *13*(5), 285-296. https://doi.org/10.1002/cpp.498

Hartley, S., Haddock, G., & Barrowclough, C. (2012). Anxiety and depression and their links with delusions and hallucinations in people with a dual diagnosis of psychosis and substance misuse: A study using data from a randomised controlled trial. *Behaviour Research and Therapy*, *50*(1), 65-71. https://doi.org/10.1016/j.brat.2011.10.007

Hayashi, N., Igarashi, Y., & Harima, H. (2021). Delusion progression process from the perspective of patients with psychoses: A descriptive study based on the primary delusion concept of Karl Jaspers. *PLoS ONE*, *16*(4). https://doi.org/10.1371/journal.pone.0250766

Hofmann, A. B., Schmid, H. M., Jabat, M., Brackmann, N., Noboa, V., Bobes, J., ... & Egger, S. T. (2022). Utility and validity of the Brief Psychiatric Rating Scale (BPRS) as a transdiagnostic scale. *Psychiatry Research*, *314*, 114659. https://doi.org/10.1016/j.psychres.2022.114659

Isham, L., Griffith, L., Boylan, A.-M., Hicks, A., Wilson, N., Byrne, R., Sheaves, B., Bentall, R. P., & Freeman, D. (2021). Understanding, treating, and renaming grandiose delusions: A qualitative study. *Psychology and Psychotherapy: Theory, Research and Practice*, *94*(1), 119-140. https://doi.org/10.1111/papt.12260

Jacobsen, P., McCrum, R. L., Gee, S., & Philpott, R. (2019). Clinical profiles of people with persecutory vs grandiose delusions who engage in psychological therapy during an acute inpatient admission. *Journal of Psychiatric Intensive Care*, *15*(2), 67-78. https://doi.org/10.20299/jpi.2019.009

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261-276. https://doi.org/10.1093/schbul/13.2.261

Kiran, C., & Chaudhury, S. (2009). Understanding delusions. *Industrial Psychiatry Journal*, *18*(1), 3-18. https://doi.org/10.4103/0972-6748.57851

Knowles, R., McCarthy-Jones, S., & Rowse, G. (2011). Grandiose delusions: A review and theoretical integration of cognitive and affective perspectives. *Clinical Psychology Review*, *31*(4), 684-696. https://doi.org/10.1016/j.cpr.2011.02.009

Knowles, R., Tai, S., Jones, S. H., Highfield, J., Morriss, R., & Bentall, R. P. (2007). Stability of self-esteem in bipolar disorder: Comparisons among remitted bipolar patients, remitted unipolar patients and healthy controls. *Bipolar Disorders*, *9*(5), 490-495. https://doi.org/10.1111/j.1399-5618.2007.00457.x

Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind & Language*, *15*(1), 184-218. https://doi.org/10.1111/1468-0017.00129

Lecomte, T., Corbière, M., & Laisné, F. (2006). Investigating self-esteem in individuals with schizophrenia: Relevance of the self-esteem rating scale-short form. *Psychiatry Research*, *143*(1), 99-108. https://doi.org/10.1016/j.psychres.2005.08.019

Lensen, S. (2023). When to pool data in a meta-analysis (and when not to)? *Fertility and Sterility*, *119*(6), 902–903. https://doi.org/10.1016/j.fertnstert.2023.03.015

Lovibond, S. H., & Lovibond, P. F. (1995). Depression anxiety stress scales. *Psychological Assessment*. https://doi.org/10.1037/t01004-000

Ma, L. L., Wang, Y. Y., Yang, Z. H., Huang, D., Weng, H., & Zeng, X. T. (2020). Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: What are they and which is better? *Military Medical Research*, *7*(1), 7-7. https://doi.org/10.1186/s40779-020-00238-8

Mark, W., & Toulopoulou, T. (2016). Psychometric properties of "community assessment of psychic experiences": review and meta-analyses. *Schizophrenia Bulletin*, *42*(1), 34-44. https://doi.org/10.1093/schbul/sbv088

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy*, *28*(6), 487-495. https://doi.org/10.1016/0005-7967(90)90135-6

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, *134*(4), 382-389. https://doi.org/10.1192/bjp.134.4.382

Moritz, S., Göritz, A. S., Gallinat, J., Schafschetzy, M., Van Quaquebeke, N., Peters, M. J., & Andreou, C. (2015). Subjective competence breeds overconfidence in errors in psychosis. A hubris account of paranoia. *Journal of Behavior Therapy and Experimental Psychiatry*, *48*, 118-124. https://doi.org/10.1016/j.jbtep.2015.02.011

Moritz, S., Veckenstedt, R., Randjbar, S., Vitzthum, F., Karow, A., & Lincoln, T. M. (2010). Course and determinants of self-esteem in people diagnosed with schizophrenia during psychiatric treatment. *Psychosis*, *2*(2), 144-153. https://doi.org/10.1080/17522430903191791

Müller, M. J., Brening, H., Gensch, C., Klinga, J., Kienzle, B., & Müller, K. M. (2005). The Calgary Depression Rating Scale for schizophrenia in a healthy control group: psychometric properties and reference values. *Journal of Affective Disorders*, *88*(1), 69-74. https://doi.org/10.1016/j.jad.2005.04.005

Nagai, S. (2015). Predictors of help-seeking behavior: Distinction between help-seeking intentions and help-seeking behavior. *Japanese Psychological Research*, *57*(4), 313-322. https://doi-org.sheffield.idm.oclc.org/10.1111/jpr.12091

National Heart, Lung, and Blood Institute (NHLBI). (2013). Study quality assessment tools: A quality assessment tool for observational cohort and cross-sectional studies. Retrieved from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Neale, J. M. (1988). Defensive functions of manic episodes. In *Delusional beliefs* (pp. 138–156). John Wiley & Sons.

O'Brien, E. J., & Epstein, S. (1988). *The multidimensional self-esteem inventory: Professional manual*. Psychological Assessment Resources.

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports*, *10*(3), 799-812. https://doi.org/10.2466/pr0.1962.10.3.799

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Peralta, V., & Cuesta, M. J. (1994). Psychometric properties of the positive and negative syndrome scale (PANSS) in schizophrenia. *Psychiatry Research*, *53*(1), 31-40. https://doi.org/10.1016/0165-1781(94)90093-0

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: Introducing the PDI (Peters et al. delusions inventory). *Schizophrenia Bulletin*, *25*(3), 553-576. https://doi.org/10.1093/oxfordjournals.schbul.a033401

Peters, E., Joseph, S., Day, S., & Garety, P. (2004). Measuring delusional ideation: the 21-item Peters et al. delusions inventory (PDI). *Schizophrenia Bulletin*, *30*(4), 1005-1022. https://doi.org/10.1093/oxfordjournals.schbul.a007116

Picardi, A., Fonzi, L., Pallagrosi, M., Gigantesco, A., & Biondi, M. (2018). Delusional themes across affective and non-affective psychoses [Original Research]. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00132

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., & Roen, K. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme*. https://doi.org/10.13140/2.1.1018.4643

Radloff, L. S. (1977). The CES-D Scale:A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385-401. https://doi.org/10.1177/014662167700100306

Robson, P. (1989). Development of a new self-report questionnaire to measure self esteem. *Psychological Medicine*, *19*, 513-518. https://doi.org/10.1017/S003329170001254X

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton Univ. Pr., 1965.

Saha, S., Lim, C. C. W., Cannon, D. L., Burton, L., Bremner, M., Cosgrove, P., Huo, Y., & J. McGrath, J. (2021). Co-morbidity between mood and anxiety disorders: A systematic review and meta-analysis. *Depression and Anxiety*, *38*(3), 286-306. https://doi.org/10.1002/da.23113

Schennach, R., Riedel, M., Obermeier, M., Seemüller, F., Jäger, M., Schmauss, M., Laux, G., Pfeiffer, H., Naber, D., Schmidt, L. G., Gaebel, W., Klosterkötter, J., Heuser, I., Maier, W., Lemke, M. R., Rüther, E., Klingberg, S., Gastpar, M., & Möller, H. J. (2015). What are depressive symptoms in acutely ill patients with schizophrenia spectrum disorder? *European Psychiatry*, *30*(1), 43-50. https://doi.org/10.1016/j.eurpsy.2014.11.001

Schlier, B., Jaya, E. S., Moritz, S., & Lincoln, T. M. (2015). The Community Assessment of Psychic Experiences measures nine clusters of psychosis-like experiences: A validation of the German version of the CAPE. *Schizophrenia Research*, *169*(1), 274-279. https://doi.org/10.1016/j.schres.2015.10.034

Sheffield, J. M., Brinen, A. P., & Freeman, D. (2021). Paranoia and grandiosity in the general population: Differential associations with putative causal factors. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.668152

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*, 747-770. https://doi.org/10.1146/annurev-psych-010418-102803

Smith, B., Fowler, D. G., Freeman, D., Bebbington, P., Bashforth, H., Garety, P., Dunn, G., & Kuipers, E. (2006). Emotion and psychosis: Links between depression, self-esteem, negative schematic beliefs and delusions and hallucinations. *Schizophrenia Research*, *86*(1), 181-188. https://doi.org/10.1016/j.schres.2006.06.018

Smith, N., Freeman, D., & Kuipers, E. (2005). Grandiose delusions: An experimental investigation of the delusion as defense. *Journal of Nervous and Mental Disease*, *193*(7), 480-487. https://doi.org/10.1097/01.nmd.0000168235.60469.cc

Stefanis, N. C., Hanssen, M., Smirnis, N. K., Avramopoulos, D. A., Evdokimidis, I. K., Stefanis, C. N., Verdoux, H., & Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine*, *32*(2), 347-358. https://doi.org/10.1017/S0033291701005141

Therman, S., Suvisaari, J., & Hultman, C. M. (2014). Dimensions of psychotic experiences among women in the general population. *International Journal of Methods in Psychiatric Research*, *23*(1), 62-68. https://doi.org/10.1002/mpr.1427

Tinakon, W., & Nahathai, W. (2012). A comparison of reliability and construct validity between the original and revised versions of the Rosenberg Self-Esteem Scale. *Psychiatry Investigation*, *9*(1), 54. https://doi.org/10.4306/pi.2012.9.1.54

Ullrich, S., Keers, R., & Coid, J. W. (2014). Delusions, anger, and serious violence: new findings from the MacArthur violence risk assessment study. *Schizophrenia Bulletin*, *40*(5), 1174-1181. https://doi.org/10.1093/schbul/sbt126

Van Putten, T., Crumpton, E., & Yale, C. (1976). Drug refusal in schizophrenia and the wish to be crazy. *Archives of General Psychiatry*, *33*(12), 1443-1446. https://doi.org/10.1001/archpsyc.1976.01770120047004

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Brazilian Journal of Psychiatry*, *35*, 416-431. https://doi.org/10.1590/1516-4446-2012-1048

Warman, D. M., & Lysaker, P. H. (2011). Delusional ideation and self-esteem in individuals with psychotic disorders. *The Journal of Nervous and Mental Disease*, *199*(1), 58-61. https://doi.org/10.1097/NMD.0b013e3182044b43

Warman, D. M., Lysaker, P. H., Luedtke, B., & Martin, J. M. (2010). Self-esteem and delusion proneness. *The Journal of Nervous and Mental Disease*, *198*(6). https://doi.org/10.1097/NMD.0b013e3181e086c5

Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2021). *The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Wilson, R. S., Yung, A. R., & Morrison, A. P. (2020). Comorbidity rates of depression and anxiety in first episode psychosis: A systematic review and meta-analysis. *Schizophrenia Research*, *216*, 322-329. https://doi.org/10.1016/j.schres.2019.11.035

World Health Organization. Division of Mental Health. (1994). *Schedules for clinical assessment in neuropsychiatry.* American Psychiatric Press

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361-370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Appendix A: PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	4-6
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	6
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	8-9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	7
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	7
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	8-10
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	10-11
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	10-11
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	11
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	11-12
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	n/a
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	n/a
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	n/a

Section and Topic	Item #	Checklist item	Location where item is reported
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	13
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	12
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	n/a
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	n/a
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	n/a
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	n/a
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	12-14
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	13
Study characteristics	17	Cite each included study and present its characteristics.	15-25
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	32, 35-36
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	15-25
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	35-40
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	n/a
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	n/a
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	n/a
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	n/a
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	37-40
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	40-43

Section and Topic	Item #	Checklist item	Location where item is reported
	23b	Discuss any limitations of the evidence included in the review.	44-45
	23c	Discuss any limitations of the review processes used.	44-45
	23d	Discuss implications of the results for practice, policy, and future research.	44-45
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	6
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	6
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	8
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	n/a
Competing interests	26	Declare any competing interests of review authors.	n/a
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	n/a

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Appendix B: Adapted NHLBI Quality Assessment for Observational Studies Tool

No.	Item	Item adequately addressed? *yes/no/p/NA/NR
1.	Was the research question or objective in this paper clearly stated?	
2.	Was the study population clearly specified and defined?	
3.	Was the participation rate of eligible persons at least 50%?	
4.	Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	
5.	Was a sample size justification, power description, or variance and effect estimates provided?	
6.	For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	
7.	Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	
8.	For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	
9.	Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	
10.	Was the exposure(s) assessed more than once over time?	
11.	Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	
12.	Were the outcome assessors blinded to the exposure status of participants?	
13.	Was loss to follow-up after baseline 20% or less?	
14.	Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	

*Yes (item adequately addressed); No (item not adequately addressed); P (item partially addressed); NR (not reported); NA (not applicable)

Section 2: Empirical Study

Cross-Sectional Design Examining the Relationship Between Delusional Beliefs and Judgments of Certainty

Abstract

Objectives

Impairments in judgements of certainty have been suggested as a potentially underlying mechanism associated with delusional beliefs. This study aimed to explore the relationships between certainty judgments, compared with both response accuracy and associated response time (RT), predicting differences for individuals who experience delusional beliefs. The study also aimed to explore whether threat vs neutral question content impacted these relationships.

Methods

A between-groups cross-sectional design was implemented. Groups included patients experiencing delusional beliefs (n=22), and mental health (n=22) and general population (n=22) controls. Participants rated their certainty in attitudes towards neutral items and future events and also answered 12 threat and 12 neutral content multiple-choice general knowledge questions, with RT and self-rated certainty also being collected.

Results

Accuracy was found to positively correlate with self-rated certainty, with a non-significant weaker correlation identified for threat questions. Participants experiencing delusional beliefs had significantly lower accuracy scores but did not significantly differ in their certainty ratings compared to controls for both correct and incorrect responses.

Significant differences were found between groups for the relationship between RT and certainty ratings. Negative relationships between certainty ratings and RTs were found for both controls but not for patients experiencing delusional beliefs, regardless of question type (threat/neutral).

Conclusions

Findings suggest certainty judgments do not appear to differ for patients experiencing delusional beliefs regarding answer accuracy, but they do appear to differ for corresponding RTs. Question type (threat/neutral) was not found to significantly impact relationships. Future research should look to replicate findings in larger clinical samples.

Keywords: delusional beliefs, certainty, confidence, reasoning, decision-making, judgments

Practitioner Points:

- Individuals are more certain about correct answers and this is not different for patients experiencing delusional beliefs.
- Individuals respond quicker when they are more certain but this is not true for patients who experience delusional beliefs, who take longer to respond regardless of certainty.
- Having threat vs neutral content in questions did not significantly impact response to general knowledge questions.
- Understanding more about how patients experiencing delusions make certainty judgments could support clinical interventions and identifying at-risk individuals, but more research in larger clinical samples is needed.

Introduction

The beliefs we hold represent the reality that we judge to be likely true and which influence our interactions and choices (Connors & Halligan, 2017). Delusional beliefs are one such type, often being experienced in the context of psychotic disorders such as schizophrenia (Picardi et al., 2018). The Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-5, American Psychiatric Association [APA], 2013) specifies delusions as different from a strongly held belief in the conviction with which the belief is held despite reasonable contradictory evidence.

Within clinical literature, delusional beliefs have been described as false beliefs (Coltheart et al., 2011; Moritz et al., 2015), held with great certainty (APA, 2013), which are not amenable to change (Kiran & Chaudhury, 2009). Hence, historically, delusions have been conceptualised as "pathological beliefs characterised by irrationality" (Bayne, 2017; Sakakibara, 2016). Scholars have pointed to philosophical conundrums raised by this definition (Bentall, 2018).

The processes involved in the formation and maintenance of delusional beliefs have been explored. Research into cognitive theories of delusions has emphasised impairments of affect (Garety & Freeman, 1999), metacognitive representation (Bronstein et al., 2019), and perception and reasoning (Langdon & Coltheart, 2000). A recent meta-analysis demonstrated that delusions are associated with belief inflexibility and, specifically, an unwarranted certainty about beliefs (Zhu et al., 2018), with the Jumping-To-Conclusions (JTC) bias and Bias Against Disconfirmatory Evidence (BADE) being suggested to be linked. A meta-analysis found the JTC bias, where individuals require less information before making a decision, to be associated with delusion proneness (Dudley et al., 2015). Similarly, BADE, a reluctance of individuals to change their minds despite contradictory evidence, has been

linked with delusions (Eisenacher & Zink, 2017). A suggested mechanism that may underlie this effect is an impaired ability, or a bias in estimating certainty (Broyd et al., 2017).

Certainty Judgments

Within the literature, the terms certainty and confidence are used interchangeably (Abelson, 1988). A limitation of the literature on certainty in delusions is the failure to consider the processes involved in certainty judgments. Importantly, since the time of James, (1890), it has been recognised that certainty judgments are rarely thought through explicitly, rather they consist of an epistemic feeling (Arango-Muñoz, 2014). Verbally expressed certainty has classically been considered an estimation of the probability of choices or beliefs being correct (Caziot & Mamassian, 2021). Therefore, asking about confidence should provide insight into a person's certainty when avowing the truth of a belief (Moritz et al., 2015). DeMarree et al. (2020) found that, if asked about their certainty in attitudes towards a wide range of topics (e.g., paper plates, likelihood of WW3), certainty was shown to be an individual difference variable (some people tend to be more or less certain about everything). Nonetheless, some types of beliefs are typically held with greater certainty than others, especially religious (Thouless, 1935) and political beliefs (Costello & Bowes, 2023)

Typically, the level of certainty people hold about their beliefs is positively correlated with the accuracy of those beliefs (Koriat, 2012). A similar relationship between strength of attitude and response time (faster reaction times corresponding to greater certainty) has been demonstrated in the social and consumer psychology literature (Tormala & Rucker, 2018). This effect has been demonstrated within general population samples using a variant of the "Who Wants to be a Millionaire?" style quiz in which people stated the certainty of their answers to simple questions (Moritz et al., 2015), see below for more details.

Response Times

An internal cue that people use to decide how certain they are about their beliefs is Response Times (RTs). Because information that is easily retrieved from memory is more likely to be correct, shorter RTs are associated with an increased certainty judgement about answers (Kiani et al., 2014). An exception to this general rule occurs when people are encouraged to trade speed for accuracy. Review papers have explored the speed-accuracy trade-off with suggestions that under time pressure, individuals increase the frequency of quicker random guesses at the sacrifice of accuracy (Heitz, 2014; Wickelgren, 1977). Despite this, studies have suggested consistent evidence for quicker RTs being associated with greater certainty judgments irrespective of response accuracy (Kelley & Lindsay, 1993), particularly when participants are not influenced to select speed over accuracy (Shaw III et al., 2001).

Certainty, Accuracy and Delusional Beliefs

A possible exception to the generally observed positive correlation between certainty and accuracy has been suggested in the case of delusion-like beliefs. These symptoms have been linked with an impairment in certainty judgments as outlined in a narrative review (Balzan, 2016), and in a large general population study where increased overconfidence in errors, regarding general knowledge questions, was associated with participants scoring higher for paranoia (Moritz et al., 2015). This was done through a task based on the game "Who Wants to be a Millionaire" whereby they asked participants to answer general knowledge questions, selecting a multiple-choice response, before rating their confidence, competence, and the question difficulty. Moritz et al. (2006) also utilised a slightly different version of a "millionaire-style" quiz within a patient population (schizophrenia diagnosis). They found patients had lower thresholds of certainty required before choosing whether to respond to optional general knowledge questions and gave more incorrect responses.

These findings suggest an impairment in the meta-cognitive processes that allow certainty judgments to be made and can be interpreted as consistent with other observations of meta-cognitive deficits in patients with psychosis such as those mentioned above (JTC, BADE) and also a reduced cognitive insight. Cognitive insight is described as a willingness toward self-reflectiveness and re-appraisal of overconfidence in misinterpretations (Beck & Warman, 2004). Clinical literature suggests individuals with schizophrenia, particularly with delusions present, have reduced cognitive insight compared to controls (Engh et al., 2010).

Despite this, there is a dearth of research investigating certainty judgment in patients with delusional beliefs (Balzan, 2016). It was therefore proposed to carry out a cross-sectional study of certainty judgments and the meta-cognitive processes involved in making them, using a three-group comparison – individuals experiencing delusional beliefs, and mental health and general population controls, to establish the specificity of deficits to delusions.

Threat vs Neutral

Research has commented on some of the potential influences of emotion on higher-level cognitions such as an increased bias toward threat identification (Blanchette & Richards, 2010). Previous research into certainty judgments has generally used neutral general knowledge questions which are not associated with threat. However, because the most common content of delusions is threat-related (Freeman, 2016), and because it is plausible that meta-cognitive processes will be especially compromised when reasoning about emotionally salient events, certainty judgments for threat-related and neutral materials were explored in line with recommendations for research on reasoning in patients experiencing delusions (Connors & Halligan, 2017).

Aims and Rationale

This report is part of a larger study with a broader aim to explore potential underlying mechanisms involved in the formation and maintenance of delusional beliefs. The parallel study (completed by fellow trainee clinical psychologist, Jess Twigg) focuses on reality sharing with others.

Due to limited research into the relationship between judgments of certainty and delusional beliefs within clinical populations, this was the focus of the present study. A similar style general knowledge quiz adapted from Moritz et al. (2015) was utilised. In the hope of building on previous research, the present study attempted to explore findings more closely related to judgments about the world by exploring judgments associated with emotionally salient topics. This was done by using threat and neutral-based content within general knowledge questions.

The specific aims of the study were to examine the relationship between judgements of certainty and 1)the speed of information retrieval, and 2)the accuracy of recalled information, in individuals experiencing delusional beliefs and controls. Additional aims were to examine whether these relationships differ for threat-related and neutral materials.

Given the potential influence of the speed-accuracy trade-off, the study focussed on RTs in the context of certainty estimates as opposed to accuracy. To further counteract the potential effect, participants were blinded to the timed aspect of the study.

Hypotheses

Based on the literature review, the following hypotheses were proposed:

Hypothesis 1. Self-reported certainty in attitudes will be higher in individuals who experience delusional beliefs compared to both controls.

In a millionaire-style general knowledge quiz:

Hypothesis 2. Accuracy will be positively correlated with self-rated certainty.

Hypothesis 3. Self-rated certainty will be higher for individuals with delusional beliefs compared to the general population and mental health controls.

Hypothesis 4. Judgments of certainty will be positively correlated with faster RT. This relationship will be stronger in both controls but less strongly correlated for individuals who experience delusional beliefs.

Hypothesis 5. For all individuals, threat-based questions will have a weaker correlation between accuracy and certainty ratings than observed for neutral questions.

Hypothesis 6. For all individuals, threat-based questions will have a weaker correlation between RT and estimates of certainty than observed for neutral questions. The difference between threat vs neutral will be strongest in individuals who experience delusional beliefs.

Methods

Design

The study utilised a cross-sectional design to test preliminary hypotheses. Data was collected using the online platform Qualtrics, across two separate phases (phase 1:screening and phase 2:experimental tasks). Data collection was done jointly with the parallel study to ease participant burden (given the overlap in screening questionnaires) and to maximise responses. See Appendix A for a full breakdown of the shared and distinct aspects of the two studies. The Strengthening of the Reporting of Observational Studies in Epidemiology (STROBE; Cuschieri, 2019) was utilised for the write-up (Appendix B). The study was not pre-registered due to time constraints.

Ethical Approval and Considerations

Ethical approval was received from National Health Service (NHS) Wales Research Ethics Committee six (ethics reference:23/WA/0271; project ID:325034; Appendix C).

Given the nature of the research, the study had the potential to evoke distress, as participation involved answering questions related to distress, experiences related to mental health, beliefs, and answering threat-based questions. This was assessed within PPI involvement to ensure acceptability and to consider any required adaptations. Participants were pre-informed about study elements, reminded of their right to withdraw, debriefed, and provided with appropriate signposting to available services where necessary. Confidentiality and its limits were discussed with participants and all participants provided informed consent to take part. Participants were provided with details for the complaints process and contact information within the information sheets (Appendix D).

There was the possibility that the clinical screening tools would highlight clinically significant responses for anxiety, depression, or delusional beliefs that the individual was unaware of/not receiving support for. Participants who requested their scores for individual scales would have been provided with this information alongside interpretation and signposting advice. No participants requested this.

Participants who completed phase 1 were invited to opt-in to a prize draw to win one of two £20 Amazon vouchers; those who completed phase 2 received an additional £10 voucher as gratuity. Participants who, opted-in to a prize draw, consented to be contacted about phase 2, or wished to receive a voucher, all provided an email address. To preserve anonymity these were stored separately in a password-protected file, deleted following the prize draw.

Unique participant ID numbers were created to anonymise participant responses and any identifiable information was stored on encrypted password-protected files separately from other collected data. Only the researchers and supervisors could access this.

Participants

To allow exploration of whether judgments of certainty differ between individuals with and without delusions, the study recruited participants into three groups: people experiencing delusional beliefs, a mental health (anxiety or depression) control, and a general population control. A range of validated measures were used to support group allocation (see measures section).

Table 1 shows the inclusion and exclusion criteria. A schizophrenia spectrum disorder diagnosis was proposed as a definitive inclusion criterion, however, based on PPI feedback this was adapted. Participants were required to be 18+ years old as special developmental issues may pertain to younger individuals.

Table 1

Participant and Specific Group Inclusion and Exclusion Criteria

Group (Name)	Inclusion	Exclusion
All groups (All)	• Capacity to consent to taking part in the study • Able to read in English • 18 years old and above	• Unable to consent to taking part in the study • Unable to read in English • Under 18 years of age
Delusional beliefs (Del)	• Schizophrenia spectrum disorder diagnosis (absence does not exclude participants) • PDI-21: endorsing 8 or more items • Any anxiety (GAD-7) and depression (PHQ-9) score • Confirmed delusions by a score of 3 or greater on the P1 subscale of the PANSS	• Bipolar and manic depressant variants of schizophrenia spectrum disorder diagnosis • PDI-21: endorsing less than 8 items • A score of 1 or 2 on P1 of the PANSS
Control: mental health (depression/ anxiety) (MH)	• No schizophrenia spectrum disorder diagnosis • PDI-21: endorsing less than 8 items • Either anxiety (GAD-7) score of 8 and above **OR** depression (PHQ-9) score of 10 and above	• Schizophrenia spectrum disorder diagnosis • PDI-21: endorsing 8 or more items • Neither anxiety: GAD-7 score of 8 and above **OR** depression: PHQ-9 score of 10 and above
Control - general population (GP)	• No schizophrenia spectrum disorder diagnosis • PDI-21: endorsing less than 8 items • Both anxiety (GAD-7) score less than 8 **AND** depression (PHQ-9) score less than 10	• Schizophrenia spectrum disorder diagnosis • PDI-21: endorsing 8 or more items • Diagnosis relating to depression or anxiety • Either anxiety (GAD-7) score of 8 and above **OR** depression (PHQ-9) score of 10 and above

Note. PDI= Peter's Delusional Inventory-21 (Peters et al., 2004), PANSS= Positive and Negative Syndromes Scale (Kay et al., 1987), GAD-7= Generalised Anxiety Disorder-7 (Spitzer et al., 2006), and PHQ-9= Patient Health Questionnaire-9 (Kroenke et al., 2001).

Sample Size Calculations

G*Power was used to calculate a priori power analysis to determine the required sample size. In a systematic review of studies of over-confidence and delusions effect sizes varied between 0.71-1.45 (Balzan, 2016), equivalent to large effect sizes (Cohen, 1988). Assuming an alpha of .05, a one-way ANOVA would have a power of 0.9 to detect a significant difference with only 10 per group (see Appendix E). Subsequently, the study aimed to recruit a minimum of 30 participants split equally across groups. Given requirements for the parallel study a total sample size of 66 was aimed for.

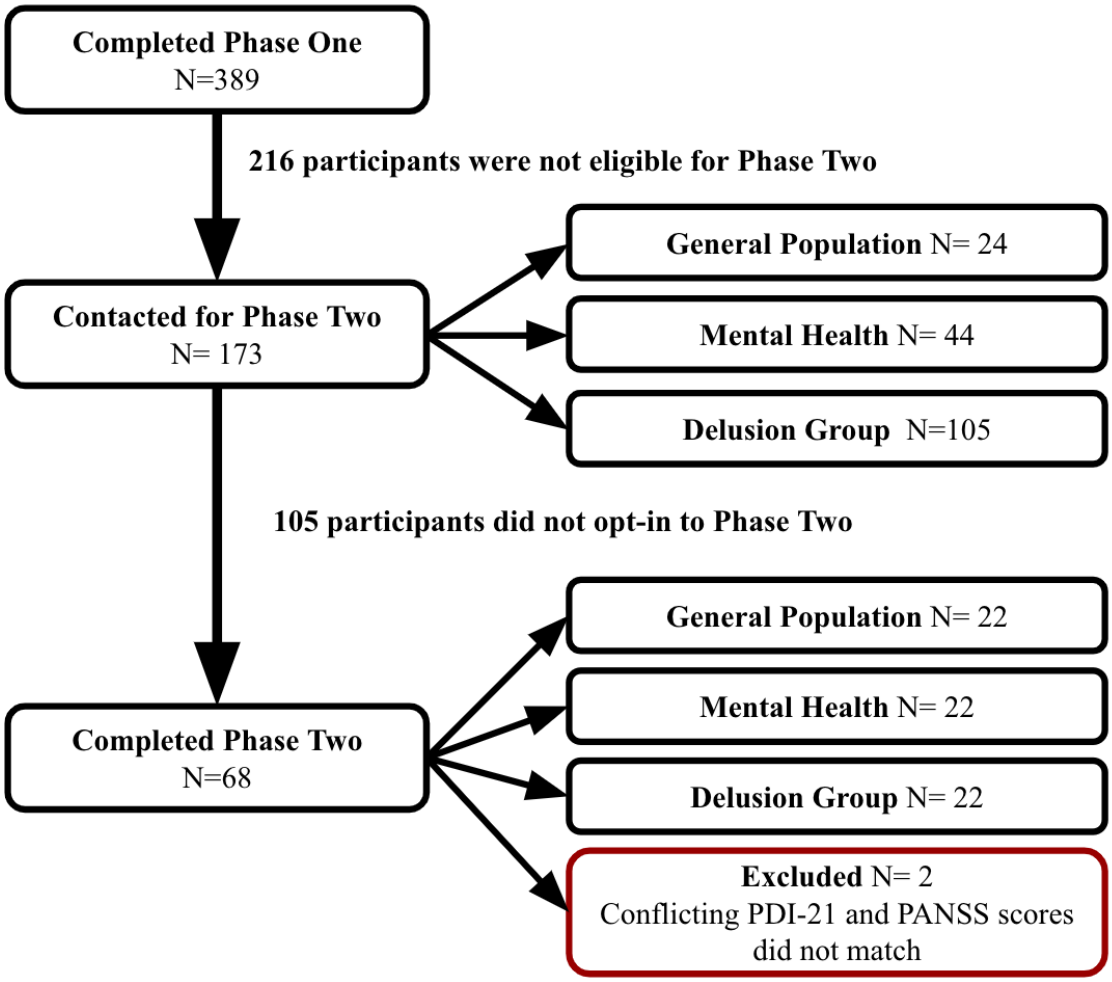
Recruitment

Recruitment was carried out by the author and fellow trainee clinical psychologist (December 2023-May 2024) through volunteer sampling using various methods. To specifically recruit for the delusions group, NHS teams in a local trust and subsequently, clinicians working with individuals with psychosis were requested to share the recruitment information with suitable individuals. Recruitment posters (Appendix F) were used to advertise online via social media (Facebook, Twitter, Reddit, and Linked In). Seven service user support organisations were also contacted via email regarding recruitment. One organisation agreed to share the recruitment poster within their monthly newsletter.

Individuals interested in taking part were able to access further information independently via Qualtrics or were invited to contact the researchers directly to further discuss. A total of 389 participants completed phase 1 and indicated they were happy to be contacted regarding phase 2. Of these, 173 were eligible and were invited to complete phase 2. Subsequently, 68 responded and successfully met with the researchers to complete phase 2, however, two individuals were excluded from analyses due to conflicting Positive and Negative Syndromes Scale (PANSS; Kay et al., 1987) scores violating the inclusion/exclusion criteria. See Figure 1 for the study flow diagram.

Figure 1

Study Flow Diagram of Participant Recruitment Process



Procedure

Phase 1- Screening

Following recruitment, participants accessed phase 1 independently on Qualtrics via the provided QR code or link. Participants were able to do this using a personal electronic device. Participants first read an information sheet before providing informed consent and contact details if they were happy to be contacted about phase 2. Participants then completed six questionnaires (see measures section).

Phase 2 – Experimental Tasks

Following phase 1, participants who consented were contacted to arrange a time to meet to complete phase 2, either online (Google Meets) or face-to-face according to participant preference. Participants followed a personalised link to access phase 2 on Qualtrics and read an information sheet before providing informed consent.

To confirm group allocation, the researchers first conducted a short clinical interview and scored individuals on the three rating scales from the PANSS (P1-delusions, P5-grandiosity, and P6-suspiciousness). To check for consensus in ratings, a second researcher also independently rated 30% (n=20) of participants on the PANSS. There was a 100% consensus between the two researchers. Depending on eligibility to continue, participants either progressed through the study or were informed that was the end of the study. Those who were asked to continue then completed a series of experimental tasks (see measures section). All participants who met with the researchers were debriefed and compensated for their time.

Measures

See Appendix H-N for the test battery order and individual questionnaires.

Phase 1

Demographic Information. Participants were asked to provide demographic information, such as gender, age, ethnicity, religion, diagnoses, and treatment.

Peter's Delusional Inventory-21 (PDI-21; Peters et al., 2004). Used as an indication of active experiences of delusional beliefs. The PDI consist of 21 items whereby individuals indicate a yes or no response (creating a total yes/no score). For *yes* items participants then rate their associated level of distress, preoccupation, and conviction, all on 5-point Likert scales. The sum of these three scales created a PDI total score. It is a reliable

and valid measure used within research as an indication of delusional ideation (Jones & Fernyhough, 2007; Peters et al., 2004). This study: Cronbach's α=.941. The cut-off score of 8 for the PDI yes/no score was selected using the mean scores referenced in Peters et al. (2004).

Generalised Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006). Used as an indication of clinically significant levels of anxiety. It is a valid and reliable tool commonly used in research on paranoid ideation (Zhu et al., 2018). This study: Cronbach's α=.866. Scores range from 0-21 with higher scores indicating a greater severity of anxiety. A cut-off score of 8 is recommended to optimise sensitivity without compromising the specificity of detection of anxiety (Plummer et al., 2016).

Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001). Used as an indication of clinically significant levels of depression. It is a valid and reliable screening tool for measuring depression (Martin et al., 2006), used in research with patients experiencing psychosis (Moritz et al., 2017). This study: Cronbach's α=.896. Scores range from 0-27, with higher scores indicating a greater severity of depression. A score of 10 indicates a moderate severity of depression with treatment being recommended (Kroenke et al., 2001).

Confidence Judgment Questionnaire (CJQ). Adapted from DeMarree et al. (2020) participants were asked about their attitude (*extremely unfavourable–extremely favourable*; 7-point scale) and certainty about their attitude (*not at all certain–extremely certain*; 7-point scale) for four neutral items (e.g., playing chess), and four possible future events (e.g., the next UK Prime Minister will be a woman). Certainty ratings were summed to create an overall CJQ certainty score. This study: Cronbach's α=.791.

Phase 2

Positive and Negative Syndromes Scale (PANSS). A brief interview based on the Structured Clinical Interview for PANSS-positive items was used to confirm the presence or absence of delusional beliefs. The PANSS is a widely used interview-based measure of

psychotic symptoms. For this study, only sub-items P1, P5 and P6 were scored relating to delusions, grandiosity, and suspiciousness/persecution, respectively. Scores range from 1(*absent*) to 7(*extreme*).

Millionaire RT Game. Adapted from the methods carried out in Moritz et al., (2015). Participants were asked 24 multiple-choice questions (12 neutral and 12 threat-based general knowledge questions) and asked to select their answer from four possible choices. Questions ranged in difficulty, were not numbered or labelled as "threat" or "neutral", and were presented one at a time in a random order to each participant.

RTs were recorded in seconds for the last click on the page before submitting, recording the response time of their selected answer. On the page following, participants were then asked to rate their certainty in their answer (*not at all certain-extremely certain*, 7-point scale) and provide an estimate of the difficulty rating on the question (*very easy-difficult*, 4-point scale).

General knowledge questions used were selected through piloting potential questions sourced from quizzes, game shows or government data (see piloting). This study: for all questions Cronbach's α=.674.

Larger Project

Given this study was completed as part of a larger project, other measures were also collected alongside those described above. See Appendix O for details.

PPI and Piloting

Questionnaire Development

To select all general knowledge questions, a larger number of questions were piloted with 19 separate individuals from a general population pool. Participants answered 24 neutral and 24 threat-based general knowledge questions with 12 of each being selected that

represented a range of difficulties, established through % of individuals getting questions correct. All questions selected were answered correctly by at least 50% of the pilot sample.

Acceptability

Due to the novel nature of some measures, phase 2 was piloted with 10 individuals. Additionally, both phases 1 & 2 were piloted with a further two individuals with lived experiences of having a MH diagnosis (depression and anxiety, schizophrenia). They reported that the methods were acceptable and they did not feel overburdened or distressed by completing the measures.

Recruitment

The study and recruitment material were discussed with clinicians working in local NHS Early Intervention Psychosis and Community Mental Health Teams. Feedback was received around the use of diagnostic-led language/inclusion criteria within recruitment materials and how this may not fit individual experiences or approaches within some services, potentially negatively impacting recruitment and representativeness. Recruitment and inclusion criteria were adapted accordingly.

Dissemination

Participants were invited to opt-in to receive a copy of the study findings. It is planned that these will be sent alongside an invite to give feedback that would hopefully guide the write-up of the study for publication. It is also planned that study findings will be disseminated to relevant services that advertised the study and requested feedback upon completion.

Statistical Analyses

Data was exported from Qualtrics into Microsoft Excel and subsequently Statistical Package for the Social Sciences (SPSS; version 29), used for all data analyses. Means,

standard deviations (SDs), confidence intervals (CIs) and distributions of categorical variables are reported throughout. Before all analyses, assumptions were tested. See Appendix P for an outline of assumption testing and findings for each analysis.

Descriptive Statistics

Group differences were assessed using Pearson's chi-square analysis (categorical data) or a one-way ANOVA (continuous data).

Inferential Statistics

Confidence Judgment Questionnaire. A one-way ANOVA was carried out to test for group differences in the total CJQ score.

Millionaire Game.

Accuracy. A total accuracy score based on the number of questions answered correctly was created for each individual, for all questions (0-24) and threat and neutral questions separately (both 0-12). Group differences in accuracy for threat vs neutral questions were explored using a two-way mixed ANOVA.

Response Time. Recorded in seconds. As RT can be, in principle, infinitely long the decision was made to remove all values 2 SD above the mean, to remove any responses that may be unrepresentative of a 'true' response (i.e., impacted by distraction, external stimuli/events). This method aligns with recommendations for dealing with RT data (Berger & Kiefer, 2021). This resulted in 31, 28, and 29 individual question RTs being removed from the delusion, MH and GP groups, respectively. The mean number of RT remaining for each participant was 22.67 and every participant had a minimum of 21 RT values following removal.

Individual RT x Certainty Correlations Coefficients. Individual RT and certainty correlations for each participant were created as suggested in Monin & Oppenheimer (2005). Pearson correlation coefficients were calculated correlating individual's RT with their

associated certainty scores for all question types together, and for threat vs neutral questions separately. As noted above, individual RTs that were greater than 2SD from the participant mean were not considered. The creation of these within-participant correlations was repeated for all responses and correct responses only. Those included for correct responses only were based on 314, 376, and 391 pairs in the Del, MH, and GP respectively. The mean number of pairs included from each participant was 16.38 and every participant had a minimum of 6 RT pairs following this removal. Correlations created for correct responses only were used within analyses[1]. Group differences in individual correlation coefficients (RTxCertainty) for threat vs neutral questions were explored using a two-way mixed ANOVA.

Creating Mean Variables. For each individual, a mean certainty score and mean RT were created for all questions combined, and threat and neutral-based questions separately. This process was repeated for correct-only and incorrect-only responses. A minimum of two values were required to calculate a mean. As a consequence, the total number of participants included in analyses using means differed from the total sample size.

RT. Mean RTs were transformed using a 1/variable transformation recommended by Ratcliff (1993), given the strong positive skew within the data and violation of assumptions. Reported means, SD and post-hoc pairwise comparisons are presented based on the untransformed values to aid interpretation. Group differences in RTs for threat vs neutral questions were explored using a two-way mixed ANOVA.

Certainty. Group differences in certainty for incorrect vs correctly answered neutral questions were explored using a two-way mixed ANOVA. This was repeated for threat questions.

[1] A sensitivity analysis was carried out and when the equivalent analyses were run on all items including incorrect items the findings did not materially change.

Certainty and Accuracy. A series of Pearson's product-moment correlations between an individual's total accuracy score and their associated mean certainty ratings were calculated for all participants for all questions, and for correct responses only, split for threat or neutral questions separately. Strengths of associations were interpreted as small=r|<.3, medium=.3<|r|<.5, and large=r|>.5 (Cohen, 1988).

An online calculator (Soper, 2014) was used to calculate if correlations significantly differed between threat vs neutral questions for all questions and questions answered correctly only.

Post-hoc pairwise comparisons were conducted where significant ANOVA results were identified. Bonferroni correction was applied to allow for multiple tests. See Appendix Q for associated SPSS outputs.

Results

Sample and Group Characteristics

In total, 66 participants completed both phases of the study and were included in the final dataset (Del n=22, MH control n=22, and GP control n=22).

The sample was predominantly female (60.6%, n=40) and the mean age across the total sample was 32.21 (SD=9.60, range=18-65 years). No significant differences were found in age, gender or religion between the three groups, however, there was a significant difference in ethnicity between groups. Summary demographics for the total sample and split by group can be found in Table 2.

Confirming the accuracy of the selection procedure, having a diagnosis, taking medication and having had therapy were significantly different between groups being more reported in the Del and MH groups compared to the GP group.

Table 2

Summary of Demographics for Each Group and Associated Group Difference Analyses

	Overall (n=66)	Del (n=22)	MH (n=22)	GP (n=22)	Group differences
Age (years)					
Mean	32.21	29.59	33.18	33.86	$F(2, 63) = 1.27$, $p = .289$
(SD)	(9.60)	(6.80)	(10.87)	(10.49)	
Range	18-65	18 - 49	21-65	25-60	
Gender					
Female	40	12	12	16	$X^2(4, N=66) = 3.84$, $p = .428$
Male	25	10	9	6	
Non-binary/ third gender	1	0	1	0	
Ethnicity					
Asian or Asian British	2	2	0	0	$X^2(2, N=66) = 28.05$, $p<.001$
Black, Black British, Caribbean or African	11	10	1	0	
Mixed or multiple ethnic groups	4	0	2	2	
White	48	10	18	20	
Other ethnic group	1	0	0	1	
Diagnosis					
Yes	31	16	15	0	$X^2(2, N=66) = 29.32$, $p<.001$
No	35	6	7	22	
Medication					
Yes	21	10	11	0	$X^2(2, N=66) = 15.51$, $p<.001$
No	45	12	11	22	
Treatment/ therapy					
Yes	31	15	14	2	$X^2(2, N=66) = 19.10$, $p<.001$
No	35	7	8	20	
Religion					
Christian	23	11	6	6	$X^2(12, N=66) = 17.52$, $p = .131$
Buddhist	1	0	0	1	
Other	2	2	0	0	
Atheism	10	1	6	3	
Agnostic	5	0	3	2	
No religion	24	7	7	10	
Prefer not to Say	1	1	0	0	

Note. SD= standard deviation

A series of one-way ANOVAs were conducted to determine if anxiety, depression and the presence of delusional beliefs significantly differed between groups. Means and SDs can be found in Table 3.

Anxiety (GAD-7). Anxiety scores significantly differed between groups *Welch's F*(2, 38.46)=43.39, $p<.001$. Anxiety scores were significantly lower in the GP group when compared with both the Del ($p<.001$) and MH ($p<.001$) groups. There was no significant difference in anxiety scores between the Del and MH groups ($p=.629$)

Depression (PHQ-9). Depression scores significantly differed between groups *Welch's F*(2, 34.65)=32.83, $p<.001$. Depression scores were significantly lower in the GP group when compared with both the Del ($p<.001$) and the MH ($p<.001$) groups. There was no significant difference in depression scores between the Del and MH groups ($p=.802$)

PDI Scores (Total and Yes/No). Both PDI total score and PDI yes/no scores were statistically significantly different between groups *Welch's F*(2, 35.68)=29.53, $p<.001$ and Welch's F(2, 39.89)=39.53, $p<.001$ respectively.

PDI total scores and yes/no scores were significantly higher in the Del group compared to the GP ($p<.001$) and MH ($p<.001$) groups. PDI total scores were significantly higher in the MH group compared to the GP group ($p=0.37$), however, PDI yes/no scores did not differ significantly between the MH and GP groups ($p=.493$).

PANSS (Items P1, P5, P6). PANSS scores were significantly different between groups for items P1 (*Welch's F*(2, 38.43)=69.88, $p<.001$), P5 (*Welch's F*(2, 34.34) =11.81, $p<.001$), and P6 (*Welch's F*(2, 38.94)=25.23, $p<.001$). PANSS item scores for all three items were significantly higher in the Del group compared to the GP and MH groups (all $p<.001$). There were no significant differences between the GP and MH groups for PANSS item P1 ($p=.513$), P5 ($p=.340$), or P6 ($p=.449$).

Table 3

Means and Standard Deviations (SD) for Screening Variables Split by Group

	Group		
	Del (n=22)	MH (n=22)	GP (n=22)
Anxiety (GAD-7)			
Mean	9.64	9.95	3.09
SD	(4.63)	(3.20)	(2.11)
Depression (PHQ-9)			
Mean	10.45	11.64	2.36
SD	(6.05)	(6.30)	(2.40)
PDI total			
Mean	110.82	30.77	16.64
SD	(56.91)	(22.07)	(13.13)
PDI yes/no			
Mean	10.95	3.55	2.86
SD	(3.84)	(1.95)	(2.01)
PANSS items			
P1			
Mean	3.68	1.27	1.14
SD	(0.95)	(0.46)	(0.35)
P5			
Mean	2.59	1.18	1.05
SD	(1.50)	(0.40)	(0.21)
P6			
Mean	3.36	1.50	1.32
SD	(1.26)	(0.51)	(0.48)

Note. PDI= Peter's Delusional Inventory-2 (Peters et al., 2004), PANSS= Positive and Negative Syndromes Scale (Kay et al., 1987), GAD-7= Generalised Anxiety Disorder-7 (Spitzer et al., 2006), and PHQ-9= Patient Health Questionnaire-9 (Kroenke et al., 2001).

Whole Sample

Accuracy and Certainty

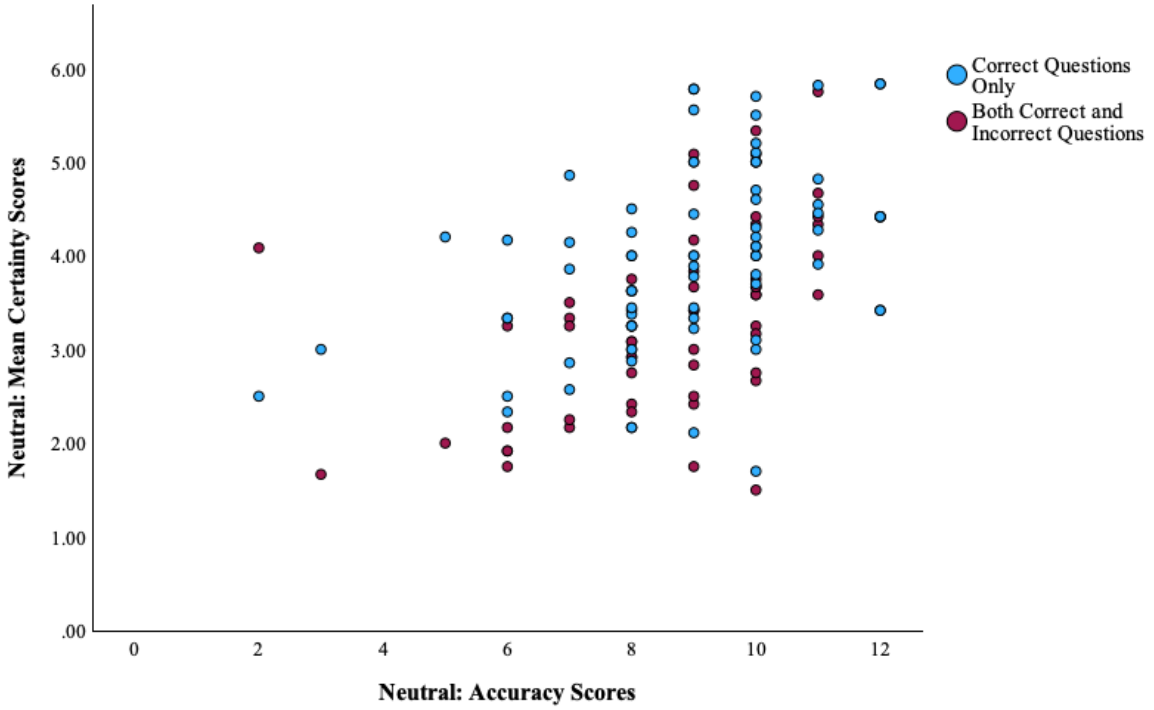
Neutral. Across the whole sample (n=66), two Pearson's product-moment correlations were run to assess the relationship between an individual's mean certainty rating and their associated accuracy for all neutral questions and neutral questions answered correctly only (see Figure 2).

Correct and Incorrect. There was a statistically significant, large positive correlation between accuracy and mean certainty for neutral questions, $r(64)=.55$, $p<.001$, with the number of questions correct explaining 30% of the variation in mean certainty scores.

Correct Only. There was a statistically significant, medium positive correlation between accuracy and mean certainty for correct neutral questions, $r(64)=.45$, $p<.001$, with the number of questions correct explaining 20% of the variation in mean certainty scores.

Figure 2

Scatterplot Illustrating the Relationship Between an Individual's Mean Certainty Rating and Their Associated Accuracy for All Neutral Questions and Neutral Questions Answered Correctly Only



Note. Each dot represents an individual and their mean certainty and the corresponding accuracy score.

Threat. Across the whole sample (n=66), two Pearson's product-moment correlations were run to assess the relationship between an individual's mean certainty ratings and their associated accuracy for all threat questions and those answered correctly only (see Figure 3).

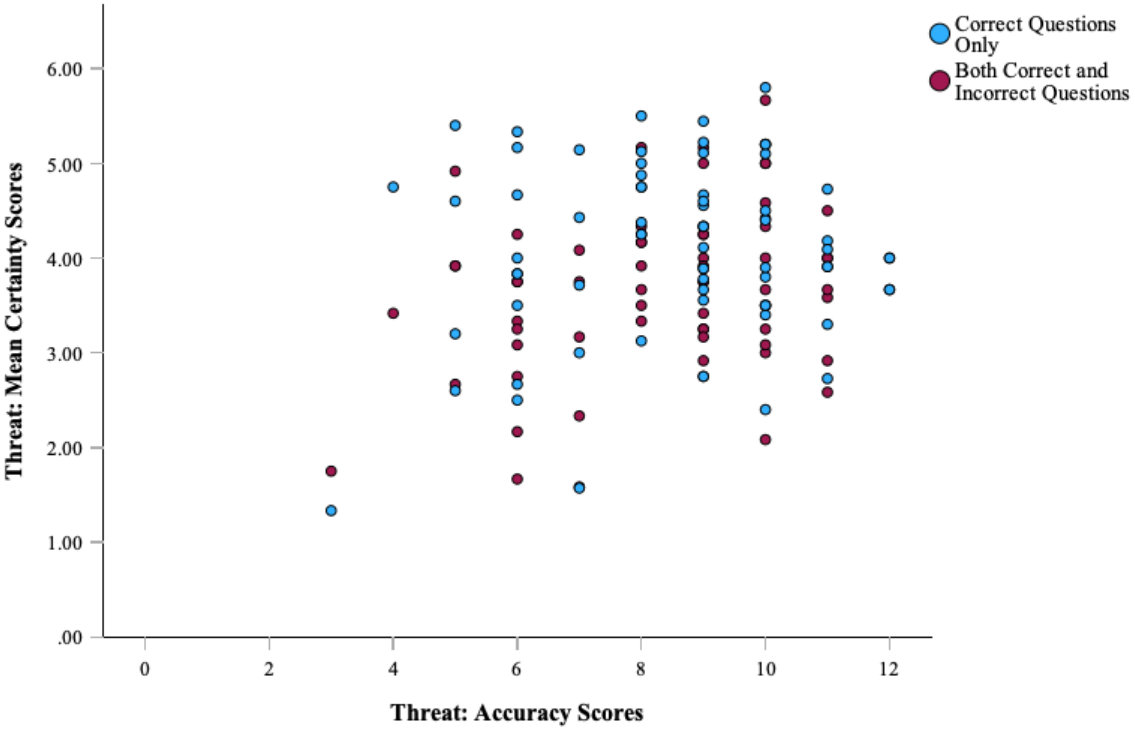
Correct and Incorrect. There was a statistically significant, small positive correlation between accuracy and mean certainty for threat-based questions, $r(64)=.28$, $p=.022$, with the number of questions correct explaining 8% of the variation in mean certainty scores.

Correct Only. There was no statistically significant correlation between accuracy and mean certainty for correct threat-based questions, $r(64)=.14$, $p=.251$, with the number of questions correct explaining 2% of the variation in mean certainty scores.

Threat vs Neutral. Accuracy and mean certainty correlations calculated above narrowly failed to significantly differ between question types (threat and neutral) when considering all questions, $z=1.86$, $p=0.063$ or when considering correctly answered questions only, $z=1.93$, $p=0.054$.

Figure 3

Scatterplot Illustrating the Relationship Between an Individual's Mean Certainty Rating and Their Associated Accuracy All Threat Questions and Threat Questions Answered Correctly Only



*N*ote. Each dot represents an individual and their mean certainty and the corresponding accuracy score

Group Differences

Confidence Judgment Questionnaire

A one-way ANOVA was used to explore group differences in total CJQ score. There were no statistically significant differences in CJQ scores between the groups (Del, MH, GP), $F(2, 63)=.16$, $p=.850$ (see Table 4). Hence, on this simple measure of confidence judgments, the judgments of participants experiencing delusions do not seem to be abnormal.

Table 4

Confidence Judgment Questionnaire: Means and Standard Deviations (SD) Split by Group

	Group		
	Del (n=22)	MH (n=22)	GP (n=22)
Mean	37	37.23	37.91
(SD)	(6.81)	(4.81)	(4.59)

Millionaire Game.

Accuracy.

Threat vs Neutral. A two-way ANOVA was used to explore differences in accuracy between groups, split for question type (threat vs neutral). See Table 5 and Figure 4.

The main effect of groups showed that there was a statistically significant difference in accuracy scores between groups $F(2, 63)=6.54$, $p=.003$, partial $\eta 2=.172$. The Del group was associated with a statistically significant mean accuracy score 1.32, 95% CI[-2.52, -0.11] lower than the MH group ($p=.027$), and 1.68, 95% CI[-2.89, -0.48] lower than the GP group (p=.003). There were no other statistically significant differences between groups. The main effect of question type showed no statistically significant difference in accuracy scores for threat vs neutral questions $F(1, 63)=2.77$, $p=.101$, partial $\eta 2=.042$. Finally, there was no statistically significant interaction between group and question type (threat neutral) on mean accuracy scores, $F(2, 63)=1.49$, $p=.233$, partial $\eta 2=.045$. Hence, participants experiencing delusional beliefs showed less accuracy overall but this effect did not depend on whether they were answering neutral or threat questions.

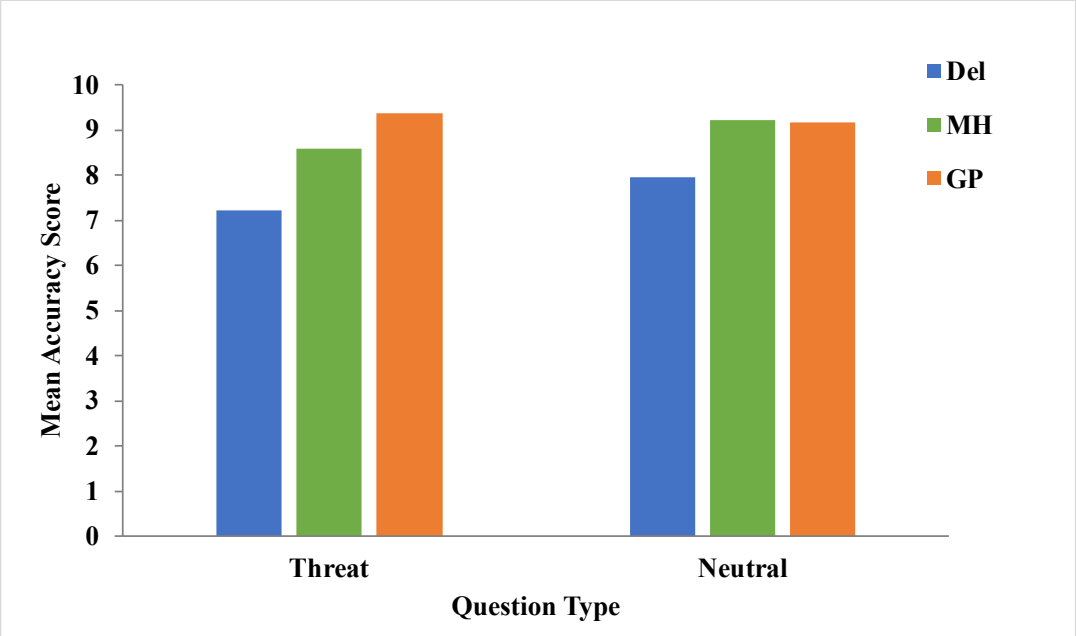
Table 5

Means and Standard Deviations (SD) for Mean Accuracy Scores Split by Group and Question Type (Threat vs Neutral)

Question type		Group		
		Del (n=22)	MH (n=22)	GP (n=22)
Threat				
	Mean	7.23	8.59	9.36
	(SD)	(2.18)	(1.87)	(1.47)
Neutral				
	Mean	7.95	9.23	9.18
	(SD)	(2.59)	(1.57)	(1.33)

Figure 4

Bar Chart Illustrating Mean Accuracy Scores (0-12) for Threat vs Neutral Questions Split by Group



Certainty.

Threat Questions: Incorrect vs Correct. A two-way ANOVA was used to look at differences in mean certainty ratings between correctly answered and incorrectly answered threat questions. See Table 6 and Figure 5.

The main effect of groups did not show a statistically significant difference in mean certainty ratings between groups $F(2, 52)=2.06$, $p=.138$, partial $\eta 2=.073$. The main effect of response type showed a statistically significant difference in certainty ratings for correct vs incorrect responses $F(1, 52)=103.26$, $p<.001$, partial $\eta 2=.665$. Certainty scores for correct answers were 1.69, 95% CI[1.36, 2.03] higher than those for incorrect answers ($p<.001$). There was no statistically significant interaction between group and response type (correct vs incorrect) on mean certainty ratings on neutral questions, $F(2, 52)=2.42$, $p=.099$, partial $\eta 2=.085$. Hence, for threat questions only, all groups rated their certainty in correct responses higher than for incorrect responses, but mean certainty ratings for individuals experiencing delusions did not differ from controls.

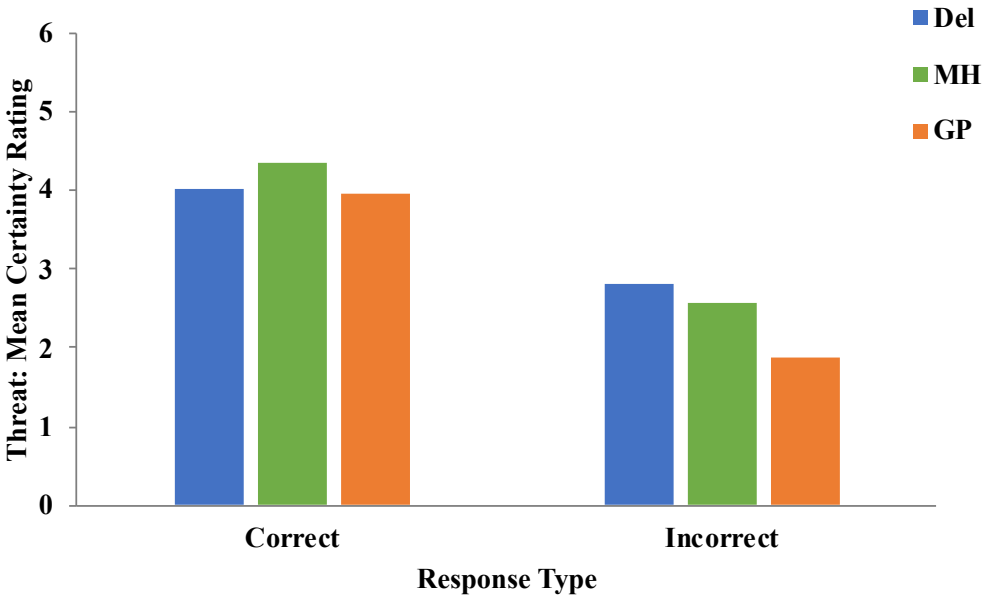
Table 6

Means and Standard Deviations (SD) for Mean Certainty Ratings (0-6) on Threat Questions Only Split by Group and Response Type (Incorrect vs Correct)

Response type	Group		
	Del (n=20)	MH (n=17)	GP (n=18)
Correct			
Mean	4.02	4.35	3.97
(SD)	(1.28)	(.74)	(.78)
Incorrect			
Mean	2.80	2.57	1.89
(SD)	(1.17)	(1.11)	(1.12)

Figure 5

Bar Chart Illustrating Mean Certainty Rating (0-6) for Threat Questions Only for Correct vs Incorrect Questions Split by Group



Neutral Questions: Incorrect vs Correct. A two-way mixed ANOVA was used to look at differences in mean certainty ratings between correctly answered and incorrectly answered neutral questions. See Table 7 and Figure 6.

The main effect of groups did not show a statistically significant difference in mean certainty ratings between groups $F(2, 48)=.249$, $p=.781$, partial $\eta 2=.010$. The main effect of response type showed a statistically significant difference in certainty ratings for correct vs incorrect responses $F(1, 48)=174.29$, $p<.001$, partial $\eta 2=.784$. Certainty scores for correct answers were 2.10, 95% CI[1.78, 2.42] higher than those for incorrect answers ($p<.001$). There was no statistically significant interaction between group and response type (correct vs incorrect) on mean certainty ratings for neutral questions, $F(2, 48)=2.08$, $p=.136$, partial $\eta 2=.080$.

Overall, based on mean certainty ratings for both neutral and threat questions, all groups rated their certainty in correct responses higher than for incorrect responses, but mean certainty ratings for individuals experiencing delusions did not differ from controls for either threat or neutral questions.

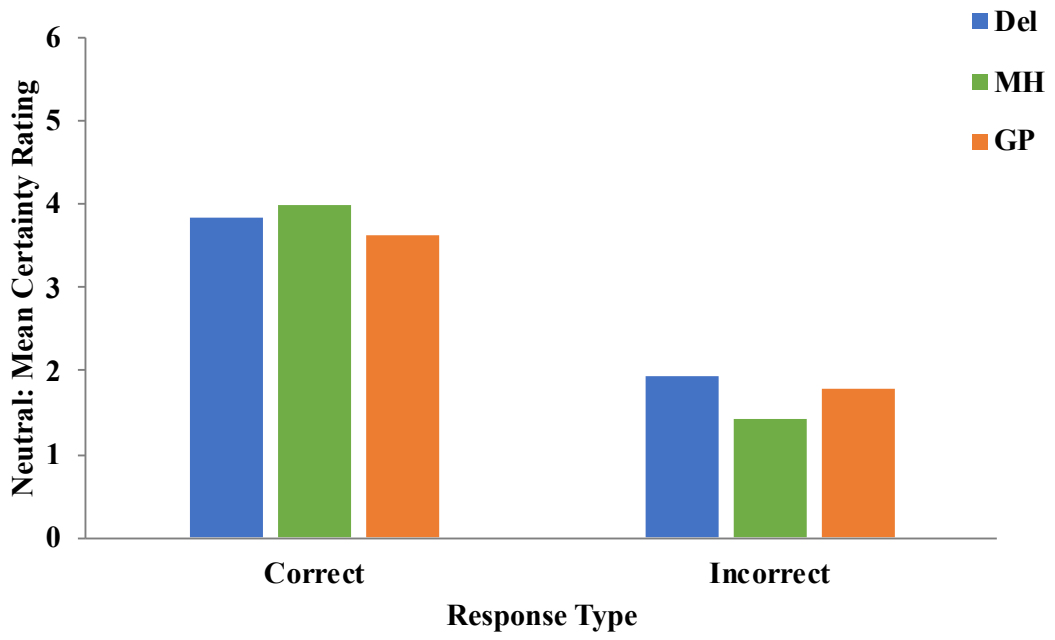
Table 7

Means and Standard Deviations (SD) for Mean Certainty Ratings (0-6) on Neutral Questions Only Split by Group and Response Type (Incorrect vs Correct)

Response type		Group		
		Del (n=20)	MH (n=17)	GP (n=18)
Correct				
	Mean	3.84	3.98	3.63
	(SD)	(1.23)	(0.91)	(.78)
Incorrect				
	Mean	1.94	1.42	1.79
	(SD)	(1.29)	(0.99)	(1.01)

Figure 6

Bar Chart Illustrating Mean Certainty Rating (0-6) for Neutral Questions Only for Correct vs Incorrect Questions Split by Group



Mean RT.

Threat Questions: Incorrect vs Correct. A two-way mixed ANOVA was used to look at differences in mean RTs between correctly and incorrectly answered threat questions. See Table 8 and Figure 7.

The main effect of groups showed a statistically significant difference in mean RTs between groups $F(2, 52)=8.72$, $p<.001$, partial $\eta 2=.251$. Based on non-transformed mean RT values, mean RTs were 8.75, 95% CI[1.71, 16.80] and 10.95, 95% CI[3.03, 18.88] longer than the MH and GP groups respectively. These were statistically significantly longer for the Del group compared to the MH ($p=.002$) and GP ($p=.004$) groups. There were no other statistically significant differences between groups. The main effect of response type did not show a statistically significant difference in mean RTs for correct vs incorrect responses $F(1, 52)=2.74$, $p=.104$, partial $\eta 2=.050$. There was no statistically significant interaction between group and response type (correct vs incorrect) on mean RTs (transformed) for threat questions, $F(2, 52)=1.78$, $p=.179$, partial $\eta 2=.064$. Hence, for threat questions participants experiencing delusions had longer RTs overall but this effect did not depend on response type and whether they were correct or incorrect.

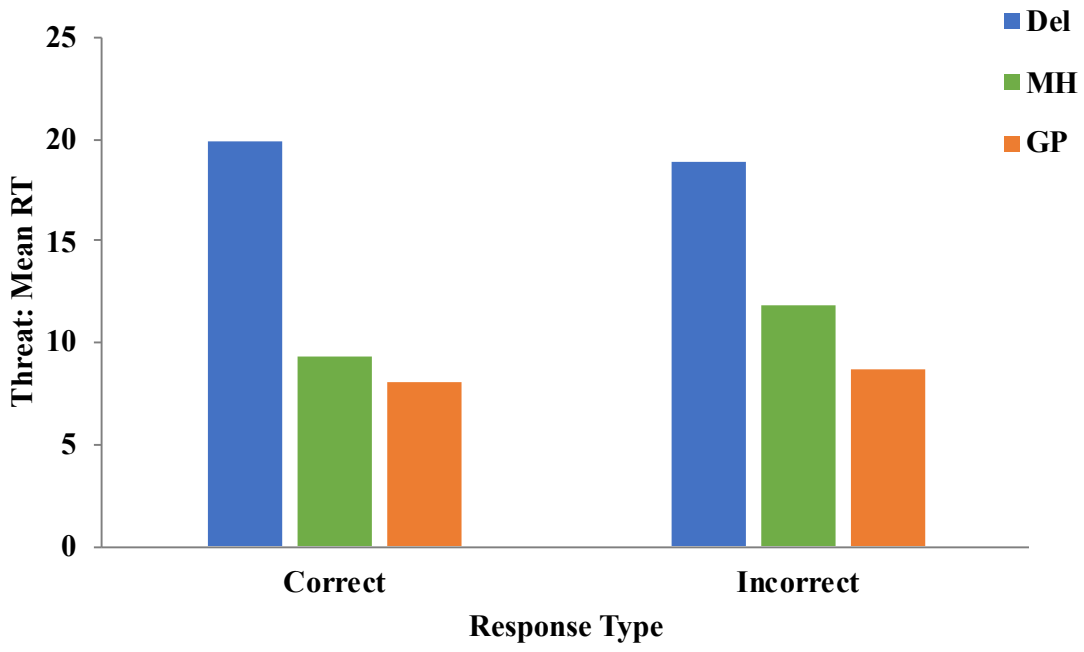
Table 8

Means and Standard Deviations (SD) for Non-Transformed Mean RTs (Seconds) on Threat Questions Only Split by Group and Response Type (Incorrect vs Correct)

Answer type	Group		
	Del (n=19)	MH (n=17)	GP (n=15)
Correct			
Mean	19.85	9.36	8.10
(SD)	(13.79)	(9.75)	(5.19)
Incorrect			
Mean	18.84	11.83	8.69
(SD)	(16.63)	(10.40)	(3.64)

Figure 7

Bar Chart Illustrating Mean RTs for Non-Transformed Threat Questions Only for Correct vs Incorrect Questions Split by Group



Neutral Questions: Incorrect vs Correct. A two-way mixed ANOVA was used to look at group differences in mean RT between correctly and incorrectly answered neutral questions. See Table 9 and Figure 8.

The main effect of groups showed a statistically significant difference in mean RTs between groups $F(2, 48)=4.31$, $p=.019$, partial $\eta 2=.152$. Based on non-transformed mean RT values, mean RTs were 6.36, 95% CI[1.58, 11.13] and 7.40, 95% CI[2.46, 12.34] longer than the MH and GP groups respectively. These were statistically significantly longer for the Del group compared to the MH ($p=.005$) and GP ($p=.002$) There were no other statistically significant differences between groups. The main effect of response type did not show a statistically significant difference in mean RTs for correct vs incorrect responses $F(1, 48)=499$, $p=.483$, partial $\eta 2=.010$. There was no statistically significant interaction between group and response type (correct vs incorrect) on mean RTs (transformed) for neutral questions, $F(2, 48)=0.28$, $p=.754$, partial $\eta 2=.012$.

Hence, a similar pattern is evident and overall, for mean RT for both threat and neutral questions, participants experiencing delusions had longer RTs compared to controls but this effect did not depend on response type and whether they were correct or incorrect.

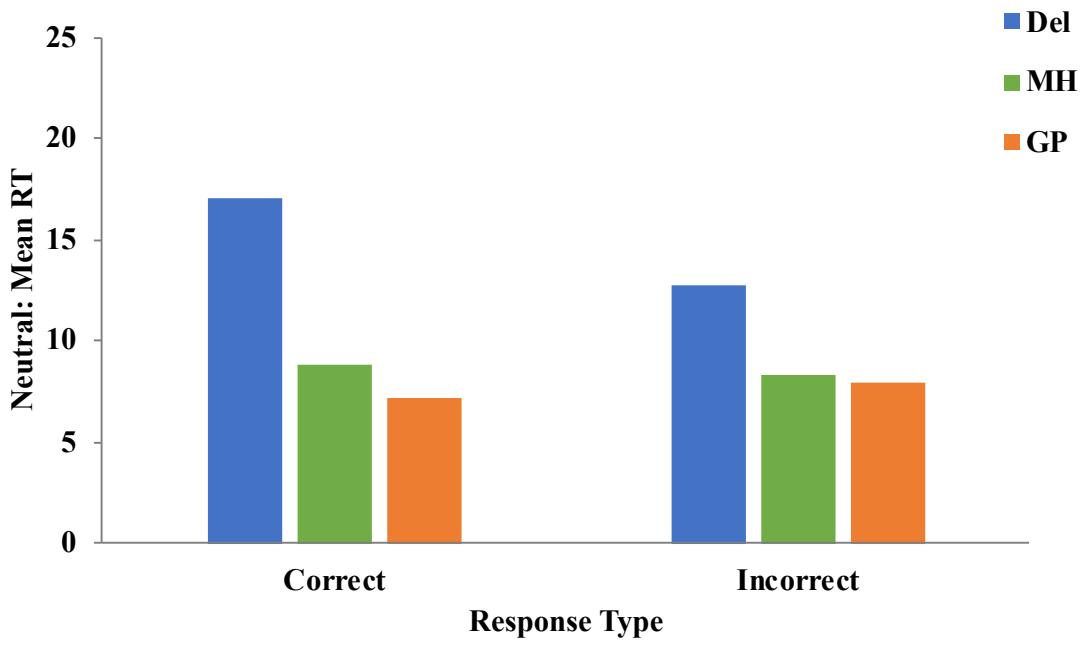
Table 9

Means and Standard Deviations (SD) for Non-Transformed Mean RTs (Seconds) on Neutral Questions Only Split by Group and Response Type (Incorrect vs Correct)

Answer type		Group		
		Del (n=19)	MH (n=17)	GP (n=15)
Correct				
	Mean	17.06	8.80	7.15
	(SD)	(13.14)	(6.23)	(2.16)
Incorrect				
	Mean	12.76	8.31	7.88
	(SD)	(7.23)	(4.03)	(2.22)

Figure 8

Bar Chart Illustrating Mean RTs (seconds) for Non-Transformed Neutral Questions Only for Correct vs Incorrect Questions Split by Group



Individual RT and Certainty Correlations

Threat vs Neutral. A two-way mixed ANOVA was used to explore group differences in the individual RT and certainty correlations created for each participant (see statistical analyses section), split for question type (threat vs neutral). The confidence intervals for the mean correlations of the groups in this analysis were used to interpret whether, for each group, the correlations differed from zero. See Table 10 and Figure 9.

The main effect of groups showed that there was a statistically significant difference in RT x Certainty correlations between groups $F(2, 63)=6.17$, $p=.004$, partial $\eta 2=.164$. The Del group demonstrated a statistically significant weaker RT x Certainty correlation by 0.37, 95% CI[0.9, 0.64] compared to the MH group ($p=.005$) and by 0.31, 95% CI[0.03, 0.58] compared to the GP group ($p=.025$). There were no other statistically significant differences between groups. The main effect of question type showed no statistically significant difference in the correlation between RT and certainty for threat vs neutral questions $F(1, 63)=1.67$, $p=.202$, partial $\eta 2=.026$. There was also no statistically significant interaction between group and question type (threat neutral) on individual correlations between RT and certainty scores, $F(2, 63)=0.28$, $p=.756$, partial $\eta 2=.009$. Hence, overall, the correlation between certainty and RT at the individual and question level was not present in the Del group compared to being present within both control groups, but this effect was not affected by whether the questions were threat-related or neutral. This suggests an abnormality in the way the Del group generate certainty judgments based on RT.

Table 10

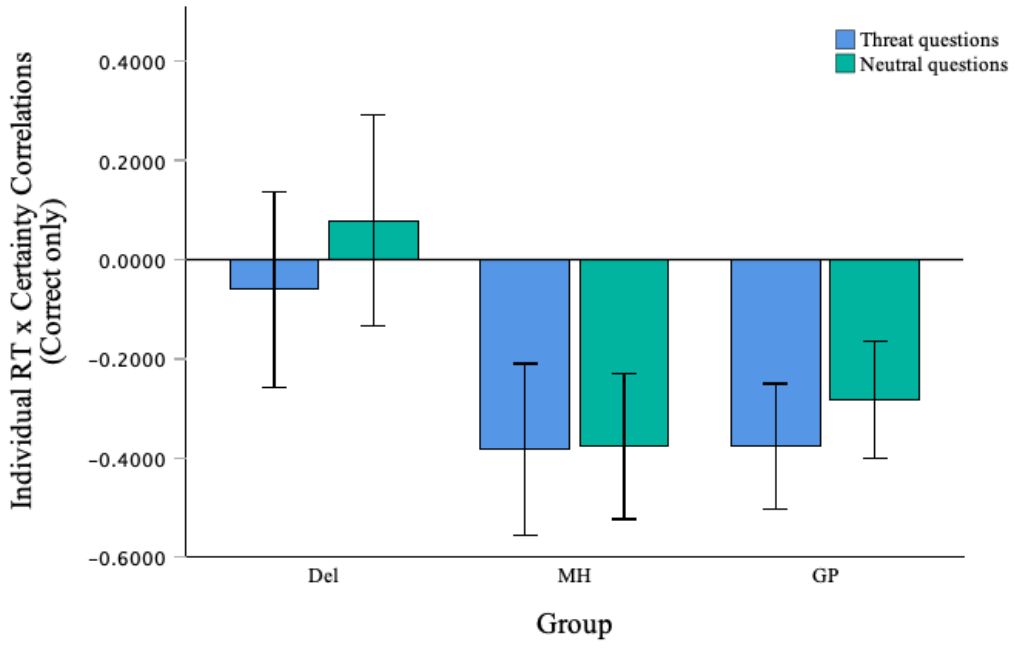
Means and Standard Deviations (SD) for Individual (RT x Certainty) Correlations for Correctly Answered Questions Only Split by Group and Question Type (Threat vs Neutral)

Question type	Group		
	Del (n=22)	MH (n=22)	GP (n=22)
Threat			
Mean	-.086	-.409	-.386
[95% CI]	[-.355, .183]	[-.598, -.221]	[-.536, -.235]
(SD)	(.606)	(.425)	(.340)
Neutral			
Mean	.015	-.397	-.297
[95% CI]	[-.211, .240]	[-.532, -.262]	[-.425, -.169]
(SD)	(.509)	(.305)	(.289)

Note. CI = 95% confidence intervals [lower bound, upper bound]

Figure 9

Bar Chart Illustrating Mean individual RT x Certainty Correlation for Correctly Answered Questions Only Threat vs Neutral Split by Group with 95% Confidence Interval Bars



Discussion

The study aimed to explore potential underlying mechanisms involved in the formation and maintenance of delusional beliefs by considering certainty judgments, specifically, the relationship between accuracy and RTs of answers with the associated self-rated certainty. It also aimed to explore whether neutral or threat-related content impacted potential differences that may exist in these relationships for individuals who experience delusions compared to controls.

Hypothesis 2 predicted that accuracy would be positively correlated with self-rated certainty. Looking at correlation findings representing the relationship between accuracy and certainty across the whole sample supported this hypothesis. Positive correlations were found suggesting a relationship whereby the more certain someone is the more accurate they are. This aligns with the findings of the level of certainty positively correlating with accuracy of beliefs (Koriat, 2012).

Looking at the potential influence of question type (threat/neutral) on this relationship, hypothesis 5 predicted weaker correlations between accuracy and certainty for threat-based vs neutral questions. Partial evidence was found to support this hypothesis. Findings showed a positive relationship between accuracy and self-rated certainty still existed for threat questions, but this was weaker than those for neutral questions and became non-significant when looking at mean certainty ratings for correctly answered threat-questions only. This suggests question type and the involvement of emotionally salient content may interrupt the relationship between certainty and accuracy. Further research is needed to explore the potential impact given analyses suggested differences to be non-significant.

Looking at findings regarding differences in certainty ratings between groups, hypotheses 1 and 3 predicted higher certainty ratings by individuals who experience delusional beliefs compared to both controls. Findings did not support this prediction.

Individuals who experience delusions do not appear to differ in their confidence ratings compared to controls on both the CJQ and the millionaire game. Both findings conflict with the widely held assumption that patients experiencing delusions hold their beliefs with unusual certainty (APA, 2013) and with the standard interpretations of the JTC, which shows patients experiencing delusions reaching a high level of certainty based on little information (Dudley et al., 2015). It might be argued that excessive certainty would only be expected for the idiosyncratic delusional beliefs of the patients but, in this context, it is worth noting that the JTC task typically employs neutral, indeed meaningless materials (Dudley et al., 2015; Huq et al., 1988).

Individuals who experience delusional beliefs had lower accuracy scores for questions compared to controls, however, findings showed certainty ratings did not differ for correct vs incorrectly answered questions. This contradicts the findings of Moritz et al. (2015) and those presented in the review by Balzan (2016) suggesting overconfidence in errors by individuals experiencing delusions.

Looking at RT, hypothesis 4 predicted certainty ratings would be positively correlated with faster RT, meaning correlations between certainty and RT would be negative (the more certain someone is the quicker they respond). It predicted that these correlations would be weaker for individuals who experience delusions. Findings provide support for this hypothesis.

When looking at the relationship between RT and certainty using individual correlation coefficients, differences for participants experiencing delusions were found. In both control groups, a stronger negative correlation was found in that the more certain an individual indicated they were then the quicker their RT was to that question. This appears to align with previous findings of quicker RTs being associated with greater certainty judgments (Kelley & Lindsay, 1993; Kiani et al., 2014). However, this was not the case for individuals

who experience delusions. For participants in the delusions group, 95% CI suggest this relationship between certainty and RTs not to be present for both threat and neutral questions. This seems to suggest there is an abnormality in how individuals experiencing delusions generate certainty judgments based on their RT. To our knowledge, no previous literature has explored the relationship between RT and certainty specifically in populations who experience delusional beliefs.

Interestingly to also note, no difference in RT associated with response type (correct/incorrect) was found in any group, suggesting no relationship between RT and accuracy. This appears to be in line with literature that suggests that RT and certainty are associated regardless of the accuracy of the response (Kelley & Lindsay, 1993). Findings showed that individuals who experience delusions had longer RT than controls, suggesting that regardless of response type they took longer to respond. The exact reason for this is unknown but possible explanations could include difficulties retrieving relevant information from memory, attentional difficulties, medication impacts (e.g., side effects of antipsychotics), or other co-occurring symptoms (i.e., depressed mood, hearing voices).

When looking at the impact of question type (threat vs neutral) on this relationship, hypothesis 6 predicted that for all individuals threat questions would result in a weaker correlation between RT and certainty compared to neutral questions. Findings did not support this. No significant difference in the relationship between RTs and certainty ratings as a result of question type was found, with individual RTxcertainty correlations for threat and neutral questions not differing significantly. This aligns with the findings above of the impact of question type on the relationship between certainty and accuracy being non-significant.

Overall, findings relating to the impact of question type (threat vs neutral) seem to suggest that having emotionally salient content (threat-related) does not significantly impact the processes associated with certainty judgments. As this was a novel approach to exploring

the impact of emotionally salient content on decision-making and judgments of certainty there is no comparable literature, however, it appears to contradict literature commenting on the impact of emotions in higher cognitive processes such as decision and judgment-making suggested by Blanchette & Richards (2010).

Strengths and Limitations

The study used both MH and GP controls which aided in being able to identify characteristics specific to individuals experiencing delusions but not present with other kinds of psychopathology. Group differences matched those desired based on allocation and groups did not differ in age, gender or religion. Two measures were used to confirm the presence of delusional beliefs increasing confidence in allocations. Steps were taken to ensure consensus in PANSS ratings supporting confidence in the reliability of scores. Changes following PPI consultation were implemented and piloting was carried out to assess acceptability and support the development of novel measures. The recruitment strategy adopted a broad approach to maximise participant recruitment of a wider variety of participants, other than just those accessing NHS services increasing generalisability.

However, there were limitations to the study. The recruitment strategy may have introduced bias with recruitment not reaching individuals who were not active online or in contact with charities or the NHS or through volunteer bias. As the majority of recruitment and collection was online, requiring an electronic device potentially resulted in a selection bias with those experiencing digital poverty being unable to take part. This is especially important given an increased risk of digital exclusion within psychosis populations (Spanakis et al., 2021). Online data collection also meant that despite being on video it was not possible to control the physical environment in which participants completed the study, with potential external stimuli possibly impacting performance. The majority of the sample was white female and groups differed significantly for ethnicity. The study also did not control for other

confounding variables such as medication, education level, employment, or attention. These points limit generalisability to wider cohorts.

Although an inclusive approach was taken with the decision to remove the requirement of a schizophrenia diagnosis in line with PPI consultation, due to smaller sample sizes it was not possible to run a sensitivity analysis regarding potential differences between those having a diagnosis or not. The study was also not pre-registered which is not in line with open science practice.

Clinical Implications

Findings highlight slower RT and differences in the relationship between certainty judgments and RT in those experiencing delusional beliefs. Better understanding whether this underpins the formation and maintenance of delusional beliefs could support the development of novel clinical interventions and with identifying at-risk populations, meaning earlier intervention which has been linked with better clinical outcomes (Fusar-Poli et al., 2017).

Additionally, understanding what may be driving the longer RT for individuals could be beneficial when planning suitable clinical interventions for individuals both in terms of adapting therapy to individual needs and increasing understanding of individual experiences.

Future Research

As this was a novel, exploratory study exploring elements of certainty judgments and decision-making in a clinical population, further research is required to replicate these findings in larger patient populations. Face-to-face approaches should be used to allow for better control of external confounding factors and other confounding cognitive factors should be controlled for (attention, other psychosis-related symptoms e.g., hearing voices). Additionally, adopting larger samples to explore these effects would support power for more complex statistical analyses and sensitivity checks for the aforementioned elements.

Additionally, more research is needed to unpick the potential impact of emotionally salient content to better replicate real-world decision-making and certainty judgments and differences in content should be included in future research. The allocation of questions as threat-related or neutral was researcher-determined and was not checked out with participants regarding the validity of allocation. It would have been beneficial within piloting to have individuals categorise statements to either neutral or threat categories or rate how much a specific question elicited an emotionally salient response to confirm the allocations. This is a specific recommendation for future research.

Another suggestion for future research would be to explore potential differences in certainty judgments specific to individual delusion belief types. Although the present study collected information on the PANSS relating to different types of delusional beliefs it was not possible to explore if any relationships discussed above differed as a result of delusion type experienced given the small sample size. Literature has suggested differences between belief types regarding associated factors (Collin et al., 2023; Grbic, 2013), potentially suggesting a non-universal mechanism within the formation and maintenance of different delusional belief types. Future research should explore differences in impairments in meta-cognitive processes such as certainty judgments split by delusion type to ascertain if identified differences are universal to all delusional beliefs or stronger for certain types. Caution will be needed for this given the often co-occurrence of multiple delusional belief types simultaneously (Collin et al., 2023; Pechey & Halligan, 2012).

Conclusion

In conclusion, the study found preliminary evidence to support the notion certainty judgments do not appear to differ for patients experiencing delusional beliefs concerning answer accuracy, but they do appear to be different in the relationship of RT and certainty judgments, not following the pattern of faster RTs for more certain responses. The study also

did not find a significant impact of emotionally salient content. As research exploring RTs and certainty judgments in this way is novel (particularly the use of emotionally salient content), further research in clinical populations adopting larger sample sizes (with more complex statistical approaches possible) is recommended to unpick this relationship.

References

Abelson, R. P. (1988). Conviction. *American Psychologist*, *43*(4), 267-275. https://doi.org/10.1037/0003-066X.43.4.267

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology*, *27*(2), 193-211. https://doi.org/10.1080/09515089.2012.732002

Balzan, R. P. (2016). Overconfidence in psychosis: The foundation of delusional conviction? *Cogent Psychology*, *3*(1). https://doi.org/10.1080/23311908.2015.1135855

Bayne, T. (2017). Delusion and the norms of rationality. In T. W. Hung & T. J. Lane (Eds.), *Rationality* (pp. 77-94). Academic Press. https://doi.org/10.1016/B978-0-12-804600-5.00005-2

Beck, A. T., & Warman, D. M. (2004). Cognitive insight: theory and assessment. In X. F. Amador & A. S. David (Eds.), *Insight and psychosis: Awareness of illness in schizophrenia and related disorders* (2nd ed., pp. 79-87). OUP Oxford.

Bentall, R. P. (2018). Delusions and other beliefs. In L. Bortolotti (Ed.), *Delusions in context* (pp. 67-96). https://doi.org/10.1007/978-3-319-97202-2

Berger, A., & Kiefer, M. (2021). Comparison of different response time outlier exclusion methods: A simulation study [Original Research]. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.675558

Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition and Emotion*, *24*(4), 561-595. https://doi.org/10.1080/02699930903132496

Bronstein, M. V., Pennycook, G., Joormann, J., Corlett, P. R., & Cannon, T. D. (2019). Dual-process theory, conflict processing, and delusional belief. *Clinical Psychology Review*, *72*, 101748. https://doi.org/10.1016/j.cpr.2019.101748

Broyd, A., Balzan, R. P., Woodward, T. S., & Allen, P. (2017). Dopamine, cognitive biases and assessment of certainty: A neurocognitive model of delusions. *Clinical Psychology Review*, *54*, 96-106. https://doi.org/10.1016/j.cpr.2017.04.006

Caziot, B., & Mamassian, P. (2021). Perceptual confidence judgments reflect self-consistency. *Journal of Vision*, *21*(12), 8-8. https://doi.org/10.1167/jov.21.12.8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M. (1997). Social ties and susceptibility to the common cold. *The Journal of the American Medical Association*, 277(24), 1940-1944. https://doi.org/10.1001/jama.1997.03540480040036

Collin, S., Rowse, G., Martinez, A. P., & Bentall, R. P. (2023). Delusions and the dilemmas of life: A systematic review and meta-analyses of the global literature on the prevalence of delusional themes in clinical groups. *Clinical Psychology Review*, *104*. https://doi.org/10.1016/j.cpr.2023.102303

Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, *62*(1), 271-298. https://doi.org/10.1146/annurev.psych.121208.131622

Connors, M. H., & Halligan, P. W. (2017). Belief and belief formation: Insights from delusions. In *Processes of believing: The acquisition, maintenance, and change in creditions* (pp. 153-165). Springer.

Costello, T. H., & Bowes, S. M. (2023). Absolute certainty and political ideology: A systematic test of curvilinearity. *Social Psychological and Personality Science*, *14*(1), 93-102. https://doi.org/10.1177/19485506211070410

Cuschieri, S. (2019). The STROBE guidelines. *Saudi Journal of Anaesthesia*, *13*(1), 31-34. https://doi.org/10.4103/sja.SJA_543_18

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2019). Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology, 32*(1), 13. https://doi.org/10.5334/irsp.198

DeMarree, K. G., Petty, R. E., Briñol, P., & Xia, J. (2020). Documenting individual differences in the propensity to hold attitudes with certainty. *Journal of Personality and Social Psychology*, *119*(6), 1239-1265. https://doi.org/10.1037/pspa0000241

Dudley, R., Taylor, P., Wickham, S., & Hutton, P. (2015). Psychosis, delusions and the "jumping to conclusions" reasoning bias: A systematic review and meta-analysis. *Schizophrenia Bulletin*, *42*(3), 652-665. https://doi.org/10.1093/schbul/sbv150

Eisenacher, S., & Zink, M. (2017). Holding on to false beliefs: The bias against disconfirmatory evidence over the course of psychosis. *Journal of Behavior Therapy and Experimental Psychiatry*, *56*, 79-89. https://doi.org/10.1016/j.jbtep.2016.08.015

Engh, J. A., Friis, S., Birkenaes, A. B., Jónsdóttir, H., Klungsøyr, O., Ringen, P. A., Simonsen, C., Vaskinn, A., Opjordsmoen, S., & Andreassen, O. A. (2010). Delusions

are associated with poor cognitive insight in schizophrenia. *Schizophrenia Bulletin*, *36*(4), 830-835. https://doi.org/10.1093/schbul/sbn193

Freeman, D. (2016). Persecutory delusions: A cognitive perspective on understanding and treatment. *The Lancet Psychiatry*, *3*(7), 685-692. https://doi.org/10.1016/S2215-0366(16)00066-3

Fusar-Poli, P., McGorry, P. D., & Kane, J. M. (2017). Improving outcomes of first-episode psychosis: An overview. *World Psychiatry*, *16*(3), 251-265. https://doi.org/10.1002/wps.20446

Garety, P. A., & Freeman, D. (1999). Cognitive approaches to delusions: A critical review of theories and evidence. *British Journal of Clinical Psychology*, *38*(2), 113-154. https://doi.org/10.1348/014466599162700

Grbic, S. (2013). *A systematic literature review of the role of self-esteem in persecutory and grandiose delusions and a grounded theory exploration of grandiose beliefs* [DClinPsy thesis, University of Sheffield]. White Rose eTheses Online. https://etheses.whiterose.ac.uk/4719/1/ethesis.pdf

Havlicek, L. L., & Peterson, N. L. (1976). Robustness of the Pearson Correlation against Violations of Assumptions. *Perceptual and Motor Skills, 43*(3), 1319-1334. https://doi.org/10.2466/pms.1976.43.3f.1319

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*, 150.

Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology. A, Human*

Experimental Psychology, *40*(4), 801-812. https://doi.org/10.1080/14640748808402300

James, W. (1890). *The principles of psychology* (Authorized ed.). Henry Holt and Company.

Jones, S. R., & Fernyhough, C. (2007). Reliability of factorial structure of the Peters et al. delusions inventory (PDI-21). *Personality and Individual Differences*, *43*(4), 647-656. https://doi.org/10.1016/j.paid.2007.01.018

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261-276. https://doi.org/10.1093/schbul/13.2.261

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*(1), 1-24. https://doi.org/10.1006/jmla.1993.1001

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*(6), 1329-1342. https://doi.org/10.1016/j.neuron.2014.12.015

Kiran, C., & Chaudhury, S. (2009). Understanding delusions. *Industrial Psychiatry Journal*, *18*(1), 3. https://doi.org/10.4103/0972-6748.57851

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80. https://doi.org/10.1037/a0025648

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind & Language*, *15*(1), 184-218. https://doi.org/10.1111/1468-0017.00129

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of Assumption Violations Revisited: A Quantitative Review of Alternatives to the One-Way Analysis of Variance F Test. *Review of Educational Research, 66*(4), 579-619. https://doi.org/10.3102/00346543066004579

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. *General Hospital Psychiatry*, *28*(1), 71-77. https://doi.org/10.1016/j.genhosppsych.2005.07.003

Martinez & Bentall., (2022) Multi-thematic Delusional Beliefs Scale. [unpublished manuscript]

Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition, 23(3),* 257-278. https://doi-org.sheffield.idm.oclc.org/10.1521/soco.2005.23.3.257

Moritz, S., Göritz, A. S., Gallinat, J., Schafschetzy, M., Van Quaquebeke, N., Peters, M. J., & Andreou, C. (2015). Subjective competence breeds overconfidence in errors in psychosis. A hubris account of paranoia. *Journal of Behavior Therapy and Experimental Psychiatry*, *48*, 118-124. https://doi.org/10.1016/j.jbtep.2015.02.011

Moritz, S., Göritz, A. S., McLean, B., Westermann, S., & Brodbeck, J. (2017). Do depressive symptoms predict paranoia or vice versa? *Journal of Behavior Therapy and Experimental Psychiatry*, *56*, 113-121. https://doi.org/10.1016/j.jbtep.2016.10.002

Moritz, S., Woodward, T. S., & Hausmann, D. (2006). Incautious reasoning as a pathogenetic factor for the development of psychotic symptoms in schizophrenia. *Schizophrenia Bulletin*, *32*(2), 327-331. https://doi.org/10.1093/schbul/sbj034

Pechey, R., & Halligan, P. (2012). Using co-occurrence to evaluate belief coherence in a large non clinical sample. *PLOS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0048446

Peters, E., Joseph, S., Day, S., & Garety, P. (2004). Measuring delusional ideation: the 21-item Peters et al. delusions inventory (PDI). *Schizophrenia Bulletin*, *30*(4), 1005-1022. https://doi.org/10.1093/oxfordjournals.schbul.a007116

Picardi, A., Fonzi, L., Pallagrosi, M., Gigantesco, A., & Biondi, M. (2018). Delusional themes across affective and non-affective psychoses [Original Research]. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00132

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510-532. https://doi.org/10.1037/0033-2909.114.3.510

Sakakibara, E. (2016). Irrationality and pathology of beliefs. *Neuroethics*, *9*(2), 147-157. https://doi.org/10.1007/s12152-016-9256-9

Shaw III, J. S., McClure, K. A., & Wilkens, C. E. (2001). Recognition instructions and recognition practice can alter the confidence–response time relationship. *Journal of Applied Psychology*, *86*(1), 93-103. https://doi.org/10.1037/0021-9010.86.1.93

Soper, D.S. (2024). *Significance of the Difference between Two Correlations Calculator* (version 4.0) [Online software]. Soper. https://www.danielsoper.com/statcalc

Spanakis, P., Heron, P., Walker, L., Crosland, S., Wadman, R., Newbronner, E., Johnston, G., Gilbody, S., & Peckham, E. (2021). Use of the internet and digital devices among people with severe mental ill health during the COVID-19 pandemic restrictions [Original Research]. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.732735

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092-1097. https://doi.org/10.1001/archinte.166.10.1092

Thouless, R. H. (1935). The tendency to certainty in religious belief. *British Journal of Psychology. General Section*, *26*(1), 16-31. https://doi.org/10.1111/j.2044-8295.1935.tb00771.x

Tormala, Z. L., & Rucker, D. D. (2018). Attitude certainty: Antecedents, consequences, and new directions. *Consumer Psychology Review*, *1*(1), 72-89. https://doi.org/10.1002/arcp.1004

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67-85. https://doi.org/10.1016/0001-6918(77)90012-9

Zhu, C., Sun, X., & So, S. H. w. (2018). Associations between belief inflexibility and dimensions of delusions: A meta-analytic review of two approaches to assessing belief flexibility. *British Journal of Clinical Psychology*, *57*(1), 59-81. https://doi.org/10.1111/bjc.12154

Appendix A: Shared and Distinct Aspects of the Projects

Researchers Daisy Fitzpatrick (DF) and Jessica Twigg (JT) completed a collaborative project. Both research topics are focused on understanding the underlying mechanism of delusional beliefs, specifically within patient populations. Due to the difficulty in recruiting participants and collecting data, paired with the time-pressured nature of Doctoral research projects, having related but different projects was planned to hopefully allow for a larger sample size to be gained.

Data collection was done jointly however specific aims, related measures and subsequent analysis differed between the two projects. The explicit similarities and differences are as follows:

Similarities

- The study design
- The sample and participant recruitment methods
- Measures for both projects were collected jointly within a single test battery (each project analysed different measures- excluding variables related to participant screening and demographics)
- Ethics application for the project was shared between researchers

Differences

- Topic area: DF assessed certainty judgments and JT assessed coalitional cognition
- Aims and hypotheses
- Data analyses
- Project write up

Aims and hypotheses for the project by DF:

Aims: To examine processes related to judgements of certainty in patients experiencing delusional beliefs compared to controls, looking specifically at certainty judgments compared with accuracy and associated Response Times (RTs). Also, to explore whether these relationships differ for threat-related and neutral materials.

Hypotheses: Predicted processes related to judgments of certainty will be different for individuals experiencing delusions with higher certainty ratings; and weaker relationships between certainty judgments and accuracy, and between judgments of certainty and RTs. The presence of threat content will impact these relationships leading to weaker correlations across all relationships compared to neutral content within questions.

Aims and hypotheses for the project by JT:

Overall Aim: This project aimed to explore a newly developing area of coalitional cognition in people with delusions. It explored one aspect of coalitional cognition, reality sharing. Two distinct areas of reality sharing were explored: (1) judgements of similarity and (2) belief sharing.

Hypotheses: People with delusions will have smaller social networks than the two control groups (H1). People with delusions will make higher plausibility ratings, especially for paranoid beliefs (H2). People with delusions will estimate that others would give higher ratings of plausibility, especially for paranoid beliefs and will not realise that other people judge these beliefs as implausible (H3). People with delusions will be less willing to discuss their beliefs with others (H4). People with delusional beliefs will not be able to judge which people are similar to them, which will impact cooperation (H5).

Appendix B: STROBE Guidelines

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	62
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	66-67
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	68-72
Objectives	3	State-specific objectives, including any prespecified hypotheses	72-73
Methods			
Study design	4	Present key elements of study design early in the paper	73
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	77-78
Participants	6	(*a*) Give the eligibility criteria and the sources and methods of selection of participants	76-79
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	79-82
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	79-82
Bias	9	Describe any efforts to address potential sources of bias	79
Study size	10	Explain how the study size was arrived at	77
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	82-85, 170
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	82-85
		(*b*) Describe any methods used to examine subgroups and interactions	82-85
		(*c*) Explain how missing data were addressed	83 170
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	n/a
		(*e*) Describe any sensitivity analyses	84 200-201

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	78, 83, 85
		(b) Give reasons for non-participation at each stage	78
		(c) Consider use of a flow diagram	78
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	86
		(b) Indicate number of participants with missing data for each variable of interest	n/a
Outcome data	15*	Report numbers of outcome events or summary measures	85-102
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	85-102
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	170-174
Discussion			
Key results	18	Summarise key results with reference to study objectives	103-106
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	106-107
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	103-109
Generalisability	21	Discuss the generalisability (external validity) of the study results	106-109
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	n/a

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Appendix C: Ethics Application and Approval Letter

Miss Fitzpatrick and Miss Twigg
Trainee Clinical Psychologist
Sheffield Health and Social Care NHS Foundation Trust
Clinical Psychology Department
Cathedral Court
Vicar Lane
S12LTN/A

Email: approvals@hra.nhs.uk
HCRW.approvals@wales.nhs.uk

10 November 2023

Dear Miss Fitzpatrick and Miss Twigg

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Belief formation in deluded and non-deluded people**
IRAS project ID:	**325034**
REC reference:	**23/WA/0271**
Sponsor	**University Of Sheffield**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The standard conditions document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **325034**. Please quote this on all correspondence.

Yours sincerely,
Anne Gell

Approvals Specialist

Email: HCRW.approvals@wales.nhs.uk

Copy to: *Professor Richard Bentall*

List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

Document	*Version*	*Date*
Copies of materials calling attention of potential participants to the research [Recruitment Poster - Delusional Belief Group]	2	19 October 2023
Copies of materials calling attention of potential participants to the research [Recruitment Poster - Mental Health Control]	2	19 October 2023
Copies of materials calling attention of potential participants to the research [Recruitment Poster - Healthy Control]	2	19 October 2023
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Sponsor Insurance]		31 July 2023
IRAS Application Form [IRAS_Form_07092023]		07 September 2023
Letter from sponsor [Sponsor Letter]		31 July 2023
Non-validated questionnaire [BET]	1	31 July 2023
Non-validated questionnaire [Demographics Q]	1	31 July 2023
Non-validated questionnaire [MDBS]	1	31 July 2023
Non-validated questionnaire [Confidence Judgement Q]	1	31 July 2023
Non-validated questionnaire [AB Game]	1	31 July 2023
Non-validated questionnaire [Debrief]	1	31 July 2023
Non-validated questionnaire [Millionaires]	1	31 July 2023
Organisation Information Document [OID]	1	19 October 2023
Other [Additional CI CV]	1	31 July 2023
Other [Additional Sponsor Letter]	1	31 July 2023
Other [Liability Certificate]	1	31 July 2023
Other [Ethical Review Response]	1	19 October 2023
Other [Ethical Review Response]	2	09 November 2023
Participant consent form [Phase two PIS and consen]	3	09 November 2023
Participant information sheet (PIS) [Phase one PIS and consent]	3	09 November 2023
Research protocol or project proposal [Study Protocol]	2	19 October 2023
Schedule of Events or SoECAT [SoECAT]	1	19 October 2023
Summary CV for Chief Investigator (CI) [CI CV]		31 July 2023
Summary CV for supervisor (student research) [Supervisor CV]		31 July 2023
Validated questionnaire [PDI]		
Validated questionnaire [GAD-7]		
Validated questionnaire [PHQ-9]		
Validated questionnaire [CSNI]		
Validated questionnaire [PANSS]		

IRAS project ID	325034

Information to support study set up

The below provides all parties with information to support the arranging and confirming of capacity and capability with participating NHS organisations in England and Wales. This is intended to be an accurate reflection of the study at the time of issue of this letter.

Types of participating NHS organisation	Expectations related to confirmation of capacity and capability	Agreement to be used	Funding arrangements	Oversight expectations	HR Good Practice Resource Pack expectations
Research activities and procedures as per the protocol and other study documents will take place at participating NHS organisations.	NHS Organisations will not be required to formally confirm capacity and capability, and research procedures may begin 35 days after provision of the local information pack, provided the following conditions are met. HRA and HCRW Approval has been issued The NHS organisation has not provided a reason as to why they cannot participateThe sponsor may start the research prior to the above deadline if the participating NHS organisation positively confirms that the research may proceed. The sponsor should now provide the local information pack to participating NHS organisations in England and/or Wales. A current list of R&D contacts is accessible at the NHS RD Forum website and these contacts MUST be used for this purpose.	An Organisation Information Document has been submitted and the sponsor is not requesting and does not expect any other agreement to be used with participating NHS organisations of this type.	A completed Schedule of Events has been provided	A Local Collaborator should be appointed at participating NHS organisations.	Where an external individual will be conducting any of the research activities that will be undertaken at this site type then they would be expected to hold a Letter of Access. This should be issued be on the basis of a Research Passport (if university employed) or an NHS to NHS confirmation of pre-engagement checks letter (if NHS employed). These should confirm Occupational Health Clearance. These should confirm standard DBS checks and appropriate barred list checks.

Other information to aid study set-up and delivery

This details any other information that may be helpful to sponsors and participating NHS organisations in England and Wales in study set-up.
The applicant has indicated that they do not intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix D: Information Sheets and Consent Forms for Phase 1 and 2

Phase 1

Title: Belief Formation in Deluded and Non-Deluded People
IRAS ID: 325034
Version: 3
Date: 09.11.2023

We would like to invite you to take part in the following research.

In this research study we will use information from you. We will only use information that we need for the research study. We will let very few people know your name or contact details, and only if they really need it for this study.

Everyone involved in this study will keep your data safe and secure. We will also follow all privacy rules.
At the end of the study we will save some of the data in case we need to check it and for future research.
We will make sure no-one can work out who you are from the reports we write.

The information that follows will tell you more about this.
Please take your time to read through it as it is important that you understand why we plan on conducting the research and what we will ask you to do if you agree to take part.

What is the purpose of this project?
Our names are Daisy Fitzpatrick and Jessica Twigg, and we are conducting this research project as part of our Doctorate of Clinical Psychology at The University of Sheffield. The University of Sheffield is the sponsor for this research.

This study aims to explore:
(1) The thought processes associated with how people form groups with others, how we cooperate within these groups, how we share beliefs with others and whether these processes are influenced by feelings of threat and trust. We will then look at whether these differ in people who have experienced delusional thoughts and those who haven't.

(2) The link between accuracy on questions and how certain someone is about an answer. To see if threat-based questions impact judgements of certainty. To look at if experiencing delusional beliefs is linked with an impacted judgement of certainty.

What will this research involve?
The study can be accessed online. There are two phases to this project.

Phase 1: phase 1 will take approximately 30-50 minutes to complete. You can access phase 1 directly after providing consent. You will then be asked to provide some personal details about you and complete some questionnaires that look at anxiety, depression, and beliefs.

We may invite you to take part in phase 2 if you meet the criteria for the study.

Phase 2: you will be asked to complete a series of questionnaires and play three short games. This phase will take approximately 55-85 minutes to complete.

The games are:
(1) AB Game: In this game you will be asked about your attitudes towards various objects e.g. soft drinks and beliefs. You will also learn about the attitudes of other people and be asked to say how similar you feel to them.

(2) Belief Exploration Task: This game will explore your beliefs, whether you think other people share them, and who you like to discuss them with.

(3) Millionaires Game: This game will involve answering some general knowledge questions by selecting a response from a choice of four answers. You will then be asked to rate your confidence in your answer and tell us how difficult you think the question was.

Benefits of taking part?
As a thank you for taking part:

If you take part in phase 1, with your consent will add you to a prize draw for a chance to win a £20 voucher.

If you take part in phase 2, you will also receive a £10 voucher as a thank you.

If you would like to be considered for phase 2 you will be asked to provide contact details in the next section.

Next we will tell you more information about how we will use your data and ask for you to consent to taking part.

How will we use information about you?
We will need to use information from you for this research project.
This information will include your:

- Name
- Email address (if provided)
- Phone number (if provided)
- Any mental health diagnoses
- Any psychiatric treatment you have received
- Any psychiatric medication you take
- Age (in years)
- Gender
- Ethnicity
- Religion
- Responses to questions

People will use this information to do the research or to check your records to make sure that the research is being done properly.
People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead. We will keep all information about you safe and secure.

Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that you took part in the study.

What are your choices about how your information is used?

- You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have.
- We need to manage your records in specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can you find out more about how your information is used?

You can find out more about how we use your information

- at www.hra.nhs.uk/information-about-patients/
- a leaflet available from https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-guidance/templates/template-wording-for-generic-information-document/
- By sending an email to one of the research team researchers: Daisy Fitzpatrick (dfitzpatrick1@sheffield.ac.uk) or Jessica Twigg (jtwigg1@sheffield.ac.uk)
- By emailing the project supervisor: Professor Richard Bentall (r.bentall@sheffield.ac.uk)
- or by emailing Amrit Sinha, Research Support Officer/Data Protection Officer: Amrit Sinha (a.sinha@sheffield.ac.uk)
- or calling 0114222 6650

Further information

Further information This research has been reviewed and approved by the NHS Wales Rec 6 (IRAS ID: 325034) and the University of Sheffield. This research will be used to write a thesis which fulfils part of the researcher's doctoral training. Please note you can contact the researchers if you would like to receive a summary of the findings.

Thank you for taking the time to read this information.

Consenting to take part

Title: Belief Formation in Deluded and Non-Deluded People

IRAS ID: 325034

Version: 3

Date: 09.11.23

Please indicate your response to the following:

	Yes	No
I confirm that I have read the information sheet (version 2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	○	○
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.	○	○
I understand that the information collected about me will be used to support other research in the future, and may be shared anonymously with other researchers.	○	○
I agree to take part in the above study.	○	○

Please write your full name:

If you want to be contacted with information about phase 2 please provide either your telephone number or email below:

Please provide a phone number in the box

Please provide an email address in the box

What if I wish to make a complaint about the way the study has been carried out?

If you wish to make a complaint or raise any concerns about this study and do not want to speak to any of the researchers, their supervisor, or the research support officer you can do this by:

- Emailing Dr Liza Monaghan (Head of Department) at l.monaghan@sheffield.ac.uk

You can also contact PALS to make a complaint. If you would like to make a complaint through PALS you can do so by:

- Writing to: Complaints Team, Sheffield Health and Social Care NHS Foundation Trust, Centre Court, Atlas Way Sheffield, S4 7QQ
- Emailing: complaints@shsc.nhs.uk or
- Calling: 0114 2718956

Further information about the PALS complaint process can be found at https://www.shsc.nhs.uk/contact-us/complaints

Phase 2

Title: Belief Formation in Deluded and Non-Deluded People
IRAS ID: 325034
Version: 3
Date: 09.11.2023

Thank you for your participation in the previous part of this study

In this research study we will use information from you. We will only use information that we need for the research study. We will let very few people know your name or contact details, and only if they really need it for this study.

Everyone involved in this study will keep your data safe and secure. We will also follow all privacy rules. At the end of the study we will save some of the data in case we need to check it and for future research. We will make sure no-one can work out who you are from the reports we write.

The information that follows will tell you more about this.
Please take your time to read through it as it is important that you understand why we plan on conducting the research and what we will ask you to do if you agree to take part.

Thank you for meeting one of our researchers online.

What is the purpose of this project?
Our names are Daisy Fitzpatrick and Jessica Twigg, and we are conducting this research project as part of our Doctorate of Clinical Psychology at The University of Sheffield. The University of Sheffield is the sponsor for this research.

This study aims to explore:
(1) The thought processes associated with how people form groups with others, how we cooperate within these groups, how we share beliefs with others and whether these processes are influenced by feelings of threat and trust. We will then look at whether these differ in people who have experienced delusional thoughts and those who haven't.

(2) The link between accuracy on questions and how certain someone is about an answer. To see if threat-based questions impact judgements of certainty. To look at if experiencing delusional beliefs is linked with an impacted judgement of certainty.

What will this research involve?
There are two phases to this project and this online questionnaire relates to phase 2.

Before you begin phase 2, you will be asked to answer some questions to confirm your eligibility to continue.

Then in phase 2 you will be asked to complete further questionnaires and play three short games. This phase will take approximately 55-85 minutes to complete.

The games are:
(1) AB Game: In this game you will be asked about your attitudes towards various objects e.g. soft drinks and beliefs. You will also learn about the attitudes of other people and be asked to say how similar you feel to them.

(2) Belief Exploration Task: This game will explore your beliefs, whether you think other people share them, and who you like to discuss them with.

(3) Millionaires Game: This game will involve answering some general knowledge questions by selecting a response from a choice of four answers. You will then be asked to rate your confidence in your answer and tell us how difficult you think the question was.

Once you have completed phase 2, the study will end, and you will be debriefed by the researcher.

Benefits of taking part?
You will receive a £10 voucher as a thank you for meeting with us today.

Next we will tell you more information about how we will use your data and ask for you to consent to taking part.

How will we use information about you?
We will need to use information from you for this research project.
This information will include your:

- Name
- Email address (if provided)
- Phone number (if provided)
- Any mental health diagnoses
- Any psychiatric treatment you have received
- Any psychiatric medication you take
- Age (in years)
- Gender
- Ethnicity
- Religion
- Responses to questions

People will use this information to do the research or to check your records to make sure that the research is being done properly.
People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead. We will keep all information about you safe and secure.

Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that you took part in the study.

What are your choices about how your information is used?

- You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have.
- We need to manage your records in specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can you find out more about how your information is used?

You can find out more about how we use your information

- at www.hra.nhs.uk/information-about-patients/
- a leaflet available from https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/data-protection-and-information-governance/gdpr-guidance/templates/template-wording-for-generic-information-document/
- By sending an email to one of the research team researchers: Daisy Fitzpatrick (dfitzpatrick1@sheffield.ac.uk) or Jessica Twigg (jtwigg1@sheffield.ac.uk)
- By emailing the project supervisor: Professor Richard Bentall (r.bentall@sheffield.ac.uk)
- or by emailing Amrit Sinha, Research Support Officer/Data Protection Officer: Amrit Sinha (a.sinha@sheffield.ac.uk)
- or calling 0114222 6650

Further information

Further information This research has been reviewed and approved by the NHS Wales Rec 6 (IRAS ID: 325034) and the University of Sheffield. This research will be used to write a thesis which fulfils part of the researcher's doctoral training. Please note you can contact the researchers if you would like to receive a summary of the findings.

Thank you for taking the time to read this information.

Consenting to take part
Title: Belief Formation in Deluded and Non-Deluded People
IRAS ID: 325034
Version: 3
Date: 09.11.23

Thank you for taking the time to read the previous information.
Please write your name and then answer the following statements.

Name:

Please indicate your response to the following:

	Yes	No
I confirm that I have read the information sheet (version 2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	○	○
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.	○	○
I understand that the information collected about me will be used to support other research in the future, and may be shared anonymously with other researchers.	○	○
I agree to take part in the above study.	○	○

What if I wish to make a complaint about the way the study has been carried out?

If you wish to make a complaint or raise any concerns about this study and do not want to speak to any of the researchers, their supervisor, or the research support officer you can do this by:

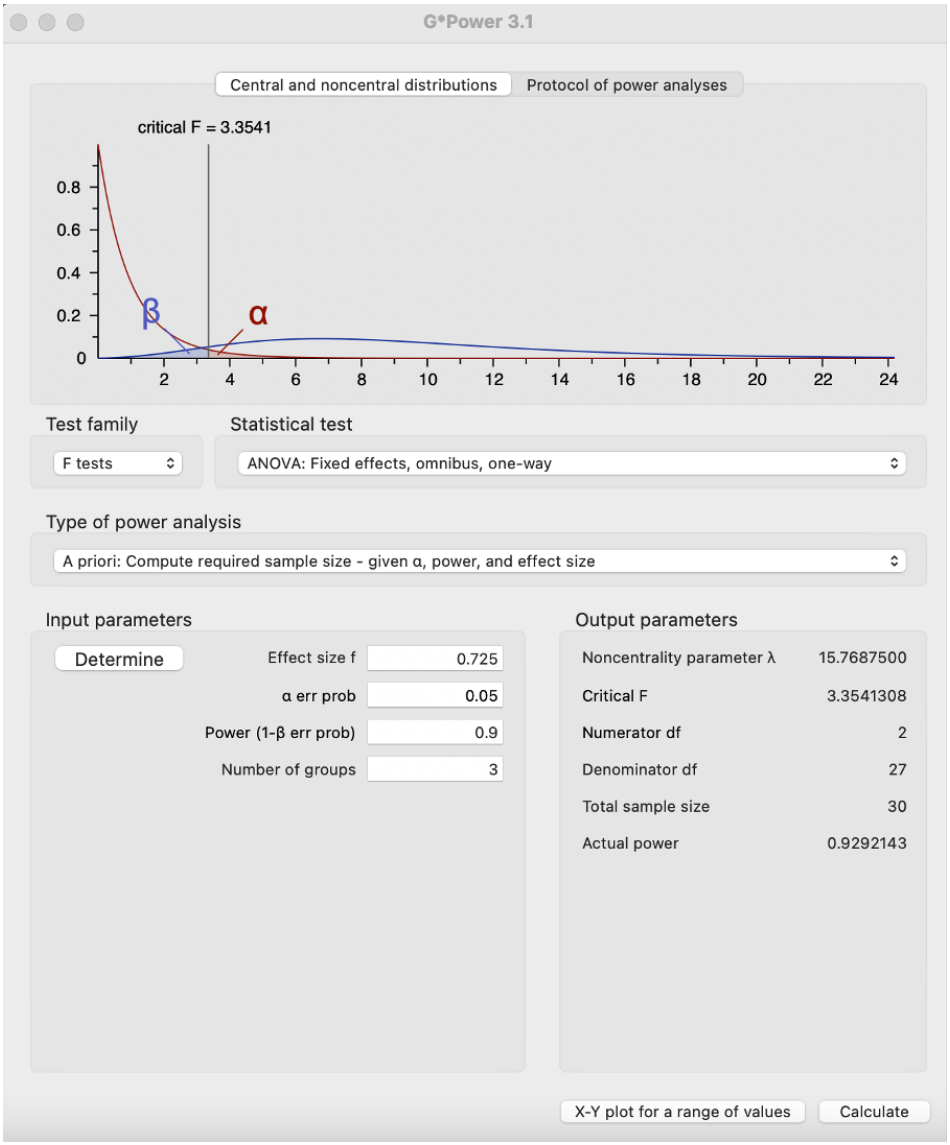
- Emailing Dr Liza Monaghan (Head of Department) at l.monaghan@sheffield.ac.uk

You can also contact PALS to make a complaint. If you would like to make a complaint through PALS you can do so by:

- Writing to: Complaints Team, Sheffield Health and Social Care NHS Foundation Trust, Centre Court, Atlas Way Sheffield, S4 7QQ
- Emailing: complaints@shsc.nhs.uk or
- Calling: 0114 2718956

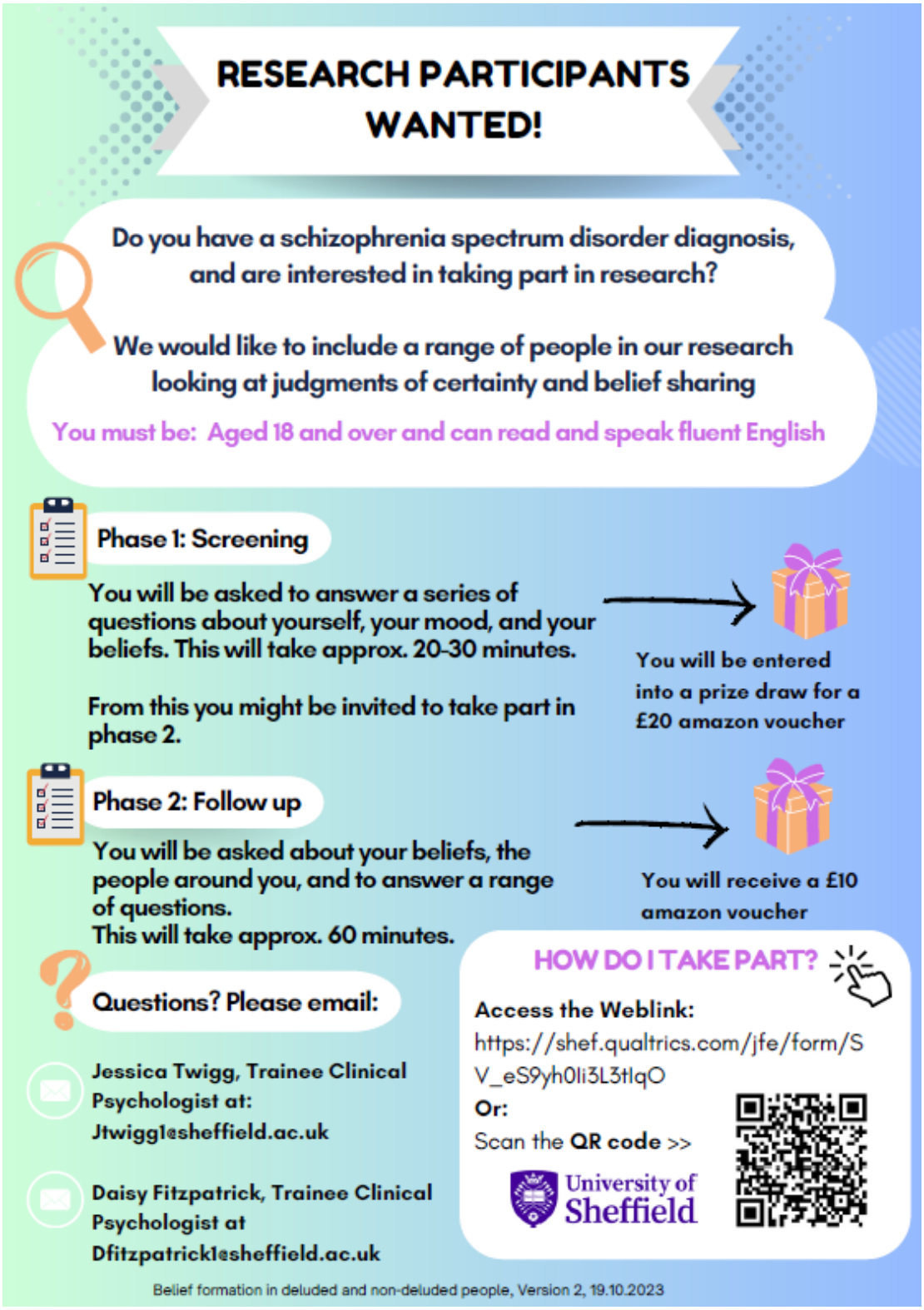
Further information about the PALS complaint process can be found at https://www.shsc.nhs.uk/contact-us/complaints

Appendix E: Priori Power Analyses: G*Power Result

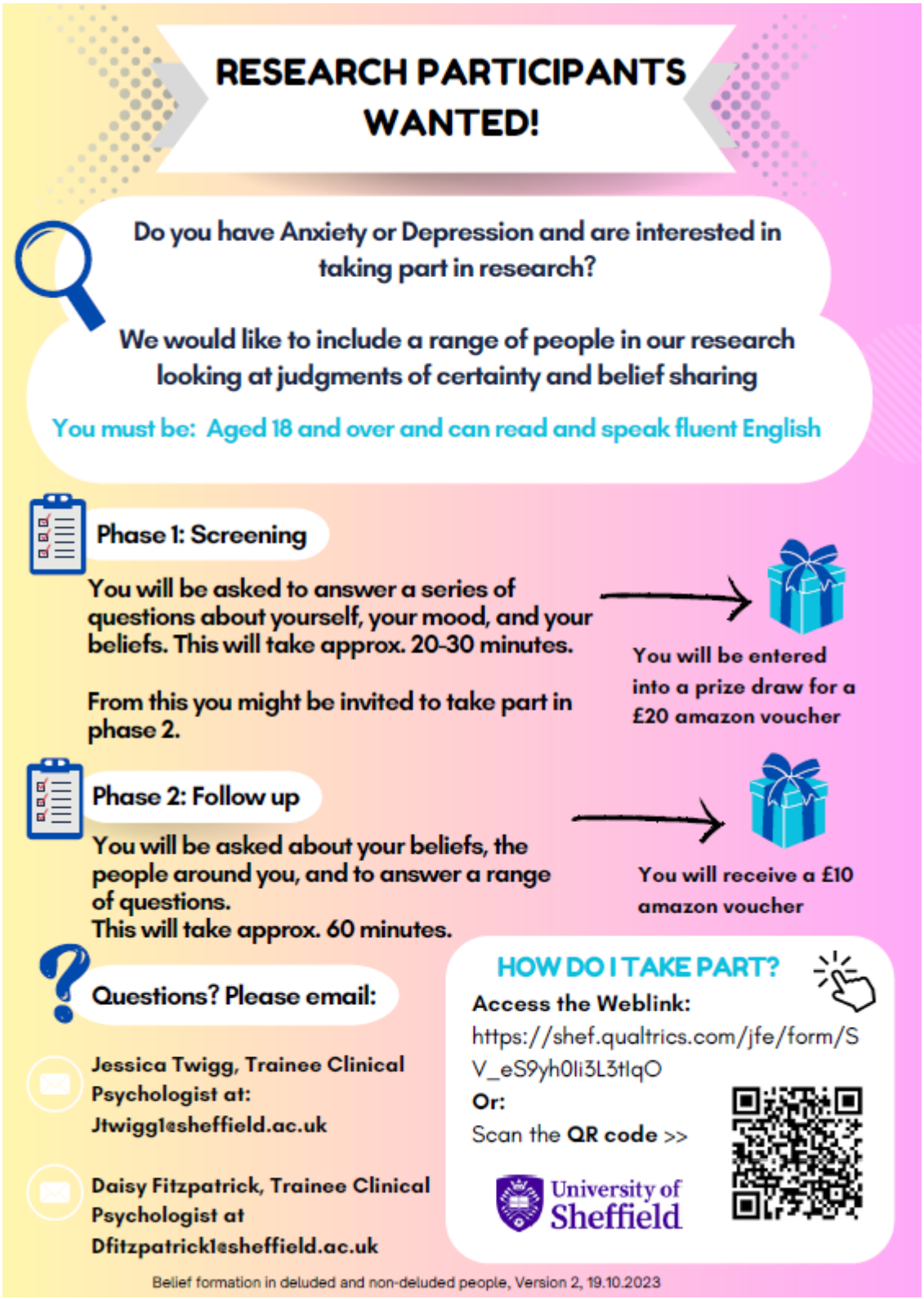


Appendix F: Recruitment Posters

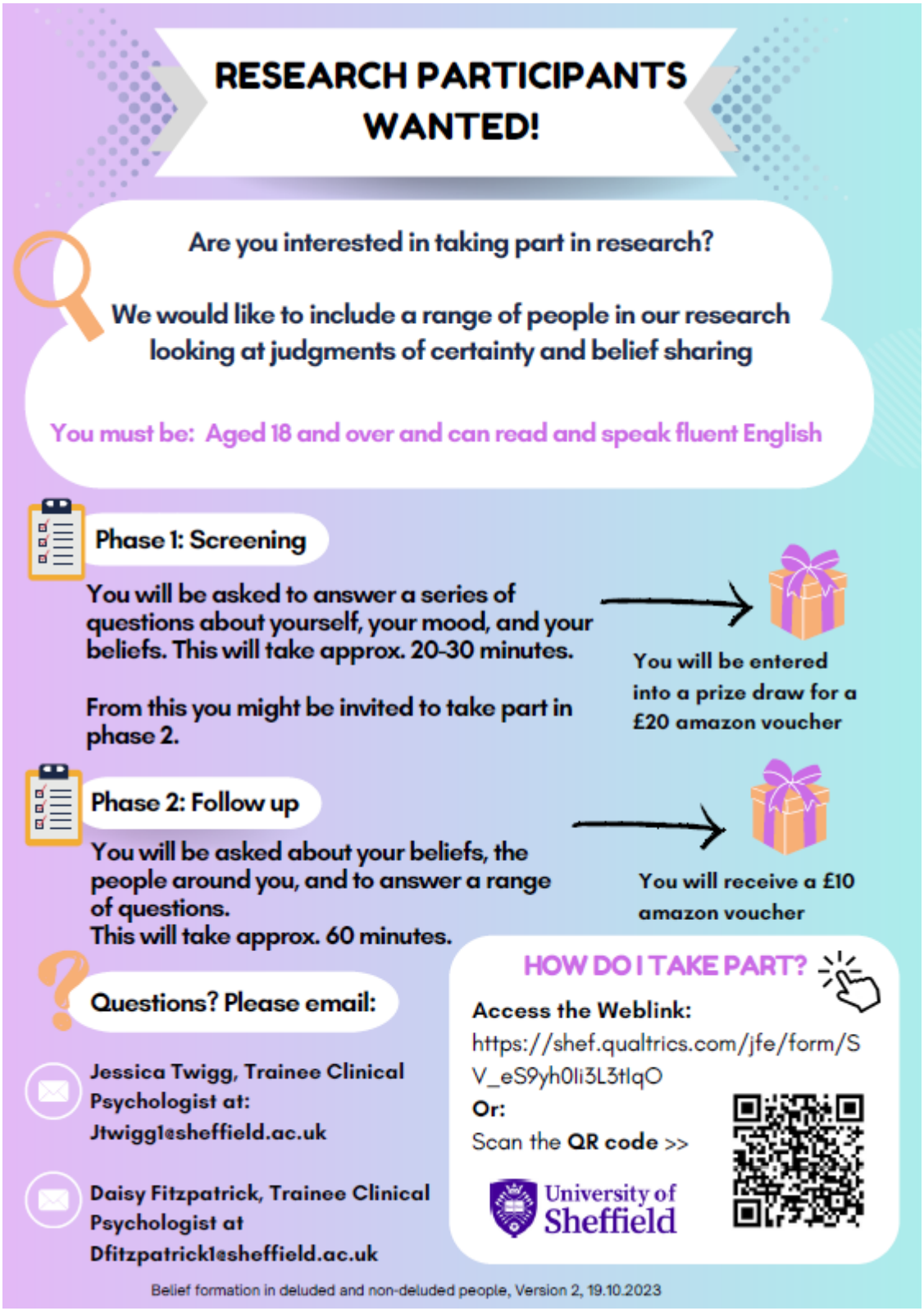
Del Group



MH Group


RESEARCH PARTICIPANTS WANTED!
Do you have Anxiety or Depression and are interested in taking part in research?
We would like to include a range of people in our research looking at judgments of certainty and belief sharing
You must be: Aged 18 and over and can read and speak fluent English
Phase 1: Screening
You will be asked to answer a series of questions about yourself, your mood, and your beliefs. This will take approx. 20-30 minutes.
From this you might be invited to take part in phase 2.
You will be entered into a prize draw for a £20 amazon voucher
Phase 2: Follow up
You will be asked about your beliefs, the people around you, and to answer a range of questions.
This will take approx. 60 minutes.
You will receive a £10 amazon voucher
Questions? Please email:
Jessica Twigg, Trainee Clinical Psychologist at:
Jtwigg1@sheffield.ac.uk
Daisy Fitzpatrick, Trainee Clinical Psychologist at
Dfitzpatrick1@sheffield.ac.uk
HOW DO I TAKE PART?
Access the Weblink:
https://shef.qualtrics.com/jfe/form/SV_eS9yh0li3L3tlqO
Or:
Scan the QR code >>
University of Sheffield
Belief formation in deluded and non-deluded people, Version 2, 19.10.2023

GP Group


RESEARCH PARTICIPANTS WANTED!
Are you interested in taking part in research?
We would like to include a range of people in our research looking at judgments of certainty and belief sharing
You must be: Aged 18 and over and can read and speak fluent English
Phase 1: Screening
You will be asked to answer a series of questions about yourself, your mood, and your beliefs. This will take approx. 20-30 minutes.
From this you might be invited to take part in phase 2.
You will be entered into a prize draw for a £20 amazon voucher
Phase 2: Follow up
You will be asked about your beliefs, the people around you, and to answer a range of questions.
This will take approx. 60 minutes.
You will receive a £10 amazon voucher
Questions? Please email:
Jessica Twigg, Trainee Clinical Psychologist at:
Jtwiggl@sheffield.ac.uk
Daisy Fitzpatrick, Trainee Clinical Psychologist at
Dfitzpatrickl@sheffield.ac.uk
HOW DO I TAKE PART?
Access the Weblink:
https://shef.qualtrics.com/jfe/form/SV_eS9yh0li3L3tlqO
Or:
Scan the QR code >>
University of Sheffield
Belief formation in deluded and non-deluded people, Version 2, 19.10.2023

Appendix G: Test Battery Order

Phase 1

1. Informed Consent

2. Screening Battery
   a. Demographics Questionnaire
   b. PDI- 21
   c. GAD-7
   d. PHQ-9
   e. CJQ
   f. MDBS

3. Screening scores analysed and next steps established:
   a. If suitable (met inclusion criteria and recruitment to the group is still ongoing) participants will be invited to complete phase 2
   b. If not suitable debrief
   c. Everyone who completes phase 1 is entered into a prize draw for a voucher (£20)

Phase 2

1. Informed Consent

2. Experimental Battery
   a. PANSS
   b. CSNI
   c. AB Game
   d. Millionaire RT Game
   e. Belief Exploration Task

3. Debrief (inc. invitation to ask for findings to be sent) and payment (£10 voucher).

Appendix H: Demographic Questionnaire

Demographic Questionnaire

We would like to ask you a few questions about yourself

Do you have any current or past mental health diagnoses?

Yes ○ No ○

What mental health diagnoses do you have?

Have you ever had or are you currently receiving psychiatric treatment? (This might include psychological therapies or pharmacological interventions).

Yes ○ No ○

What treatment have you recieved?

Do you currently take any psychiatric medication?

Yes No

Please list here:

How old are you (in years)?

How do you identify in terms of your gender?

- ◯ Male
- ◯ Female
- ◯ Non-binary / third gender
- ◯ Prefer not to say
- ◯ Other not listed above

Appendix I: Peter's Delusional Inventory-21 (PDI-21)

Appendix

P.D.I.-21

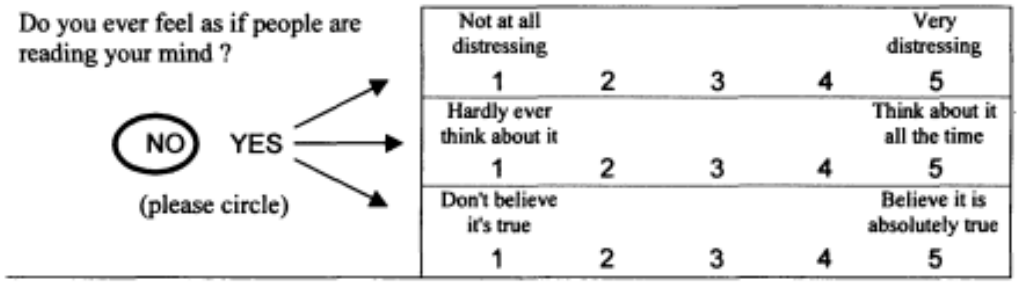
This questionnaire is designed to measure beliefs and vivid mental experiences. We believe that they are much more common than has previously been supposed, and that most people have had some such experiences during their lives. Please answer the following questions as honestly as you can. There are no right or wrong answers, and there are no trick questions.

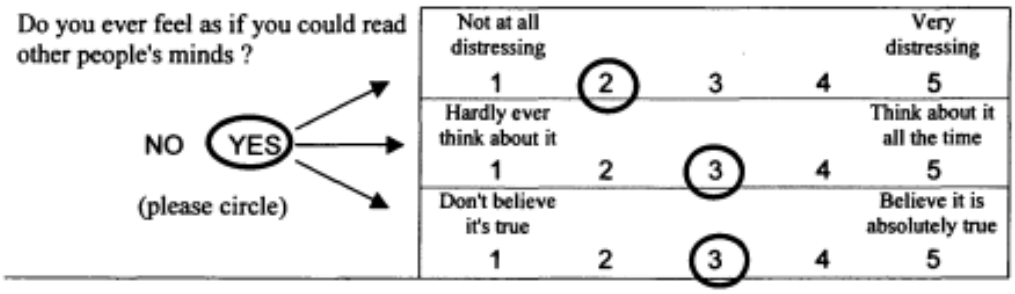
Please note that we are NOT interested in experiences people may have had when under the influence of drugs.

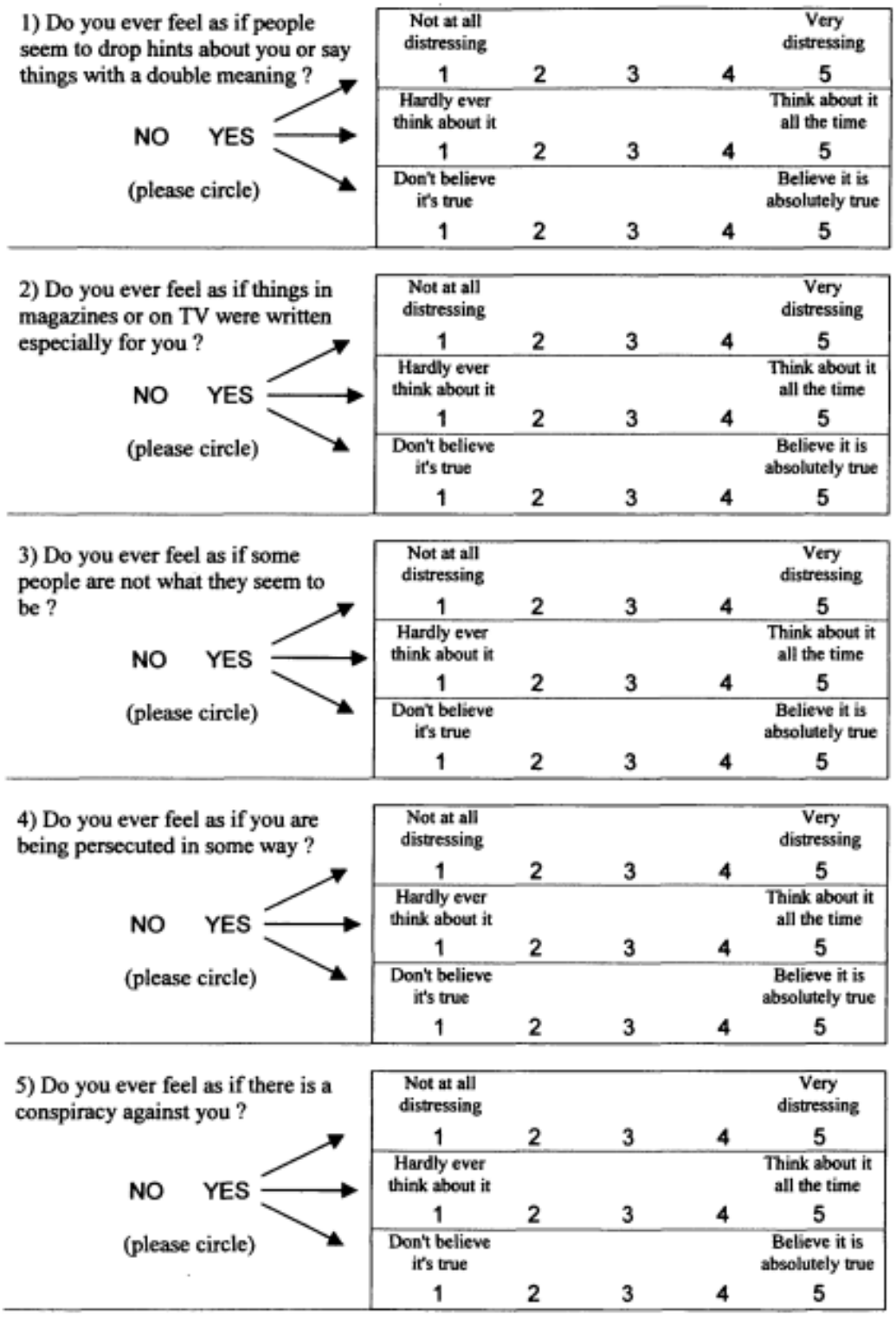
IT IS IMPORTANT THAT YOU ANSWER *ALL* QUESTIONS.

For the questions you answer YES to, we are interested in:
(a) how distressing these beliefs or experiences are
(b) how often you think about them; and
(c) how true you believe them to be.
On the right hand side of the page we would like you to circle the number which corresponds most closely to how distressing this belief is, how often you think about it, and how much you believe that it is true.
If you answer NO please move on to the next question.

Example






1) Do you ever feel as if people seem to drop hints about you or say things with a double meaning ?
NO YES
(please circle)
Not at all distressing
Very distressing
1 2 3 4 5
Hardly ever think about it
Think about it all the time
1 2 3 4 5
Don't believe it's true
Believe it is absolutely true
1 2 3 4 5
2) Do you ever feel as if things in magazines or on TV were written especially for you ?
NO YES
(please circle)
Not at all distressing
Very distressing
1 2 3 4 5
Hardly ever think about it
Think about it all the time
1 2 3 4 5
Don't believe it's true
Believe it is absolutely true
1 2 3 4 5
3) Do you ever feel as if some people are not what they seem to be ?
NO YES
(please circle)
Not at all distressing
Very distressing
1 2 3 4 5
Hardly ever think about it
Think about it all the time
1 2 3 4 5
Don't believe it's true
Believe it is absolutely true
1 2 3 4 5
4) Do you ever feel as if you are being persecuted in some way ?
NO YES
(please circle)
Not at all distressing
Very distressing
1 2 3 4 5
Hardly ever think about it
Think about it all the time
1 2 3 4 5
Don't believe it's true
Believe it is absolutely true
1 2 3 4 5
5) Do you ever feel as if there is a conspiracy against you ?
NO YES
(please circle)
Not at all distressing
Very distressing
1 2 3 4 5
Hardly ever think about it
Think about it all the time
1 2 3 4 5
Don't believe it's true
Believe it is absolutely true
1 2 3 4 5

6) Do you ever feel as if you are, or destined to be someone very important ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

7) Do you ever feel that you are a very special or unusual person ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

8) Do you ever feel that you are especially close to God ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

9) Do you ever think people can communicate telepathically ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

10) Do you ever feel as if electrical devices such as computers can influence the way you think ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

11) Do you ever feel as if you have been chosen by God in some way ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

12) Do you believe in the power of witchcraft, voodoo or the occult ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

13) Are you often worried that your partner may be unfaithful ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

14) Do you ever feel that you have sinned more than the average person ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

15) Do you ever feel that people look at you oddly because of your appearance ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

16) Do you ever feel as if you had no thoughts in your head at all ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

17) Do you ever feel as if the world is about to end ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

18) Do your thoughts ever feel alien to you in some way ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

19) Have your thoughts ever been so vivid that you were worried other people would hear them ?

NO YES

(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

20) Do you ever feel as if your own thoughts were being echoed back to you ?

NO YES

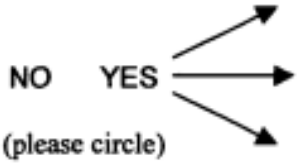
(please circle)

Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

21) Do you ever feel as if you are a robot or zombie without a will of your own ?

NO YES

(please circle)



Not at all distressing				Very distressing
1	2	3	4	5
Hardly ever think about it				Think about it all the time
1	2	3	4	5
Don't believe it's true				Believe it is absolutely true
1	2	3	4	5

Appendix J: Generalised Anxiety Disorder-7 (GAD-7)

Generalised Anxiety Disorder 7-item (GAD-7) scale

Name: _____ Date of Birth: _____

GP _____ Date Completed: _____

Over the last 2 weeks, how often have you been bothered by the following problems?	**Not at all**	**Several days**	**Over half the days**	**Nearly every day**
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it's hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3
Add the score for each column		+	+	+

Total Score (*add your column scores*) = _____

If you checked off any problems, how difficult have these made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all _____
Somewhat difficult _____
Very difficult _____
Extremely difficult _____

Appendix K: Patient Health Questionnaire-9 (PHQ-9)

Patient Health Questionnaire (PHQ-9)

Name: _____ Date: _____

Over the last 2 weeks, how often have you been bothered by any of the following problems?	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself – or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3

For office coding: Total Score _____ = _____ + _____ + _____

Total Score _____

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

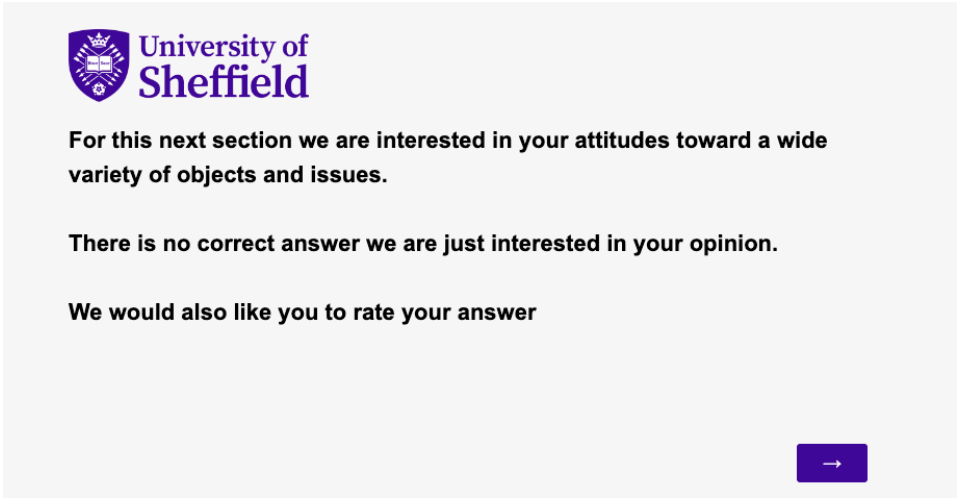
☐ Not difficult at all ☐ Somewhat difficult ☐ Very difficult ☐ Extremely difficult

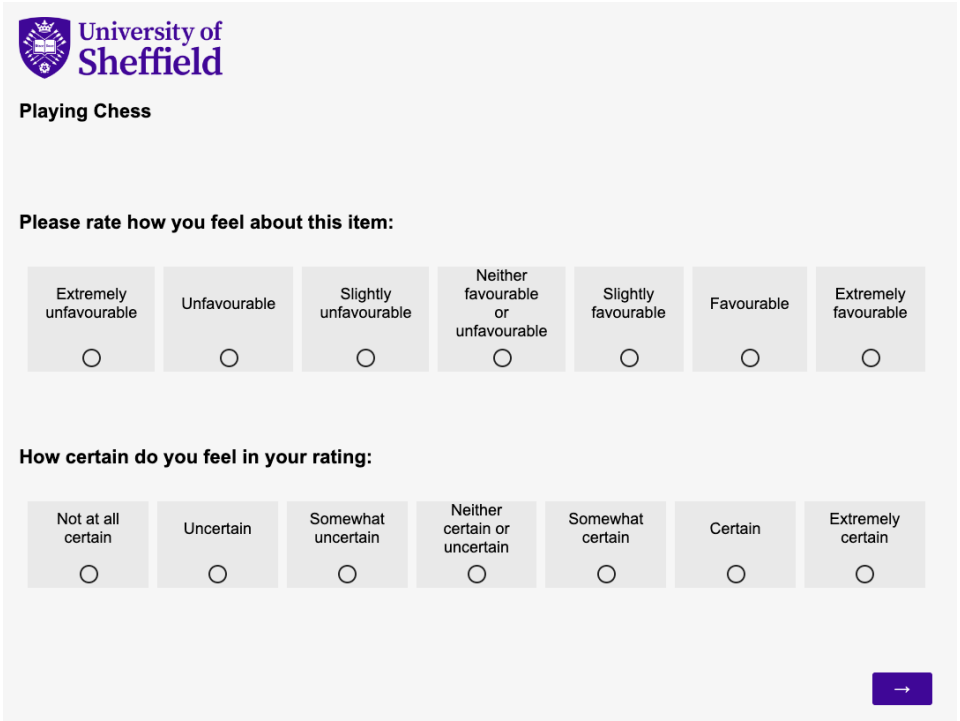
Appendix L: Confidence Judgment Questionnaire (CJQ).

Neutral items: playing chess; rugby; taxes; paper plates.

Possible future events: The next UK Prime Minister will be a woman; There will be a manned Mars mission by 2026; There will be a manned Mars mission by 2026; There will be an effective cure for lung cancer by 2040.

Instructions and Example Item Presentation





Appendix M: Positive and Negative Syndromes Scale (PANSS) Scoring

PANSS Scoring Sheet Used

Patient ID		Date		Time	
Brief notes					

P1 Delusions Beliefs which are unfounded, unrealistic, and idiosyncratic. Basis for rating thought content expressed in the interview and its influence on social relations and behaviour.	
1 Absent - Definition does not apply	
2 Minimal - Questionable pathology; may be at the upper extreme of normal limits.	
3 Mild - Presence of one or two delusions which are vague, uncrystallized, and not tenaciously held. Delusions do not interfere with thinking, social relations, or behaviour.	
4 Moderate - Presence of either a kaleidoscopic array of poorly formed, unstable delusions or of a few well-formed delusions that occasionally interfere with thinking, social relations, or behaviour.	
5 Moderate severe - Presence of numerous well-formed delusions that are tenaciously held and occasionally interfere with thinking, social relations, or behaviour.	
6 Severe - Presence of a stable set of delusions which are crystallised, possibly systematised, tenaciously held, and clearly interfere with thinking, social relations, and behaviour.	
7 Extreme - Presence of a stable set of delusions which are either highly systematised or very numerous, and which dominate major facets of the patient's life. This frequently results in inappropriate and irresponsible action, which may even jeopardise the safety of the patient or others.	

P5. Grandiosity Exaggerated self-opinion and unrealistic convictions of superiority, including delusions of extraordinary abilities, wealth, knowledge, fame, power, and moral righteousness. Basis for rating: thought content expressed in the interview and its influence on behaviour.	
1 Absent - Definition does not apply	

2 Minimal - Questionable pathology; may be at the upper extreme of normal limits.	
3 Mild - Some expansiveness or boastfulness is evident, but without clear-cut grandiose delusions.	
4 Moderate - Feels distinctly and unrealistically superior to others. Some poorly formed delusions about special status or abilities may be present but are not acted upon.	
5 Moderate severe - Clear-cut delusions concerning remarkable abilities, status, or power are expressed and influence attitude but not behaviour.	
6 Severe - Clear-cut delusions of remarkable superiority involving more than one parameter (wealth, knowledge, fame, etc.) are expressed, notably influence interactions, and may be acted upon.	
7 Extreme - Thinking, interactions, and behaviour are dominated by multiple delusions of amazing ability, wealth, knowledge, fame, power, and/or moral stature; which may take on a bizarre quality.	
P6. Suspiciousness/persecution Unrealistic or exaggerated ideas of persecution, as reflected in guardedness, a distrustful attitude, suspicious hypervigilance, or frank delusions that others mean one harm. Basis for rating: thought content expressed in the interview and its influence on behaviour.	
1 Absent - Definition does not apply	
2 Minimal - Questionable pathology; may be at the upper extreme of normal limits.	
3 Mild - Presents a guarded or even openly distrustful attitude, but thoughts, interactions, and behaviour are minimally affected.	
4 Moderate - Distrustfulness is clearly evident and intrudes on the interview and or behaviour, but there is no evidence of persecutory delusions. Alternatively, there may be indication of loosely formed persecutory delusions, but these do not seem to affect the patient's attitude or interpersonal relations.	
5 Moderate severe - Patient shows marked distrustfulness, leading to major disruption of interpersonal relations, or else there are clear-cut persecutory delusions that have limited impact on interpersonal relations and behaviour.	
6 Severe - Clear-cut pervasive delusions of persecution which may be systematised and significantly interfere in interpersonal relations.	
7 Extreme - A network of systematised persecutory delusions dominates the patient's thinking, social relations, and behaviour.	

Interview Questions Asked

Patient ID		Date		Time:	

Brief notes	

Section 1: Build Rapport
(Instruction to interviewer: Allow at least 5 minutes for a non-directive phase serving to establish rapport in the context of an overview before proceeding to the specific questions listed below.)
- *Introduction and instructions*
Section 2: Delusions and Unusual Thought Content
1. Have things been going well for you?
2. Has anything been bothering you lately?
3. Can you tell me something about your thoughts on life and its purpose?
4. Do you follow a particular philosophy (any special rules, teachings, or religious doctrine)?
5. Some people tell me they believe in the Devil; what do you think? **IF YES:** Can you tell me more about this?
6. Can you read other people's minds? **IF YES:** How does that work?
7. Can others read your mind? **IF YES:**

How can they do that? Is there any reason that someone would want to read your mind?
8. Who controls your thoughts?
Section 3: Suspiciousness/Persecution and Poor Impulse Control
9. How do you spend your time these days?
10. Do you prefer to be alone?
11. Do you join in activities with others? **IF YES** Tell me about it. **IF NO** Why not? ... Are you afraid of people, or do you dislike them? **IF YES** Can you explain?
12. Do you have many friends? **IF YES** Close friends? **IF NO** Why not? **IF NO** Just a few? **IF NO** Any? .... Why? **IF YES** Why just a few friends?
13. Do you feel that you can trust most people? **IF NO** Why not?

14. Are there some people in particular who you don't trust? **IF YES** Can you tell me who they are? Do 15 **IF NO and yes prev (can trust most people)** Skip 15 **IF NO and no to prev (can't trust most people)** Do 15
15. Why don't you trust people (or name specific person)? **IF DONT KNOW OR DONT WANT TO SAY** Do you have a good reason not to trust…? Is there something that ..(they).. did to you? Perhaps something that ..(they).. might do to you now? **IF YES** Can you explain to me?
16. Do you get along well with others? **IF NO** What's the problem?
17. Do you have a quick temper?
18. Do you like most people? **IF NO** Why not?
19. Are there perhaps some people who don't like you? **IF YES** For what reason?

20. Do others talk about you behind your back?
IF YES
What do they say about you?
Why?

21. Does anyone ever spy on you or plot against you?

22. Do you sometimes feel in danger?
IF YES
Would you say that your life is in danger?
Is someone thinking of harming you or even perhaps thinking of killing you?
Have you gone to the police for help?
Do you sometimes take matters into your own hands or take action against those who might harm you?
IF YES
What have you done?

Section 4: "Guilt Feelings" (G3) and "Grandiosity" (P5)

23. If you were to compare yourself to the average person, how would you come out: a little better, maybe a little worse, or about the same?

"BETTER,"	**"ABOUT THE SAME,"**	**"WORSE,"**
Better in what ways?	Are you special in some ways? **IF YES** In what ways?	Worse in what ways? Just how do you feel about yourself?

24. Would you consider yourself gifted?

25. Do you have talents or abilities that most people don't have?

IF YES Please explain.
26. Do you have any special powers? **IF YES** What are these? Where do these powers come from?
27. Do you have extrasensory perception (ESP)?
28. Are you very wealthy? **IF YES:** Please explain
29. Can you be considered to be very bright? **IF YES** Why would you say so?
30. Would you describe yourself as famous?
31. Would some people recognize you from TV, radio, or the newspaper? **IF YES,** Can you tell me about it?
32. Are you a religious person? **IF YES** Are you close to God?

IF YES
Did God assign you some special role or purpose?
Can you be one of God's messengers or angels?

IF YES	**IF NO:**
What special powers do you have as God's messenger (angel)?	Do you perhaps consider yourself to be God?

33. Do you have some special mission in life?
IF YES,
What is your mission?
Who assigned you to that mission?

34. Did you ever do something wrong — something you feel bad or guilty about?

IF YES
Just how much does that bother you now?
Do you feel that you deserve punishment for that?

IF YES
What kind of punishment would you deserve?
Have you at times thought of punishing yourself?

IF YES
Have you ever acted on those thoughts of punishing yourself?

Appendix N: Millionaire RT Game Questions and Display

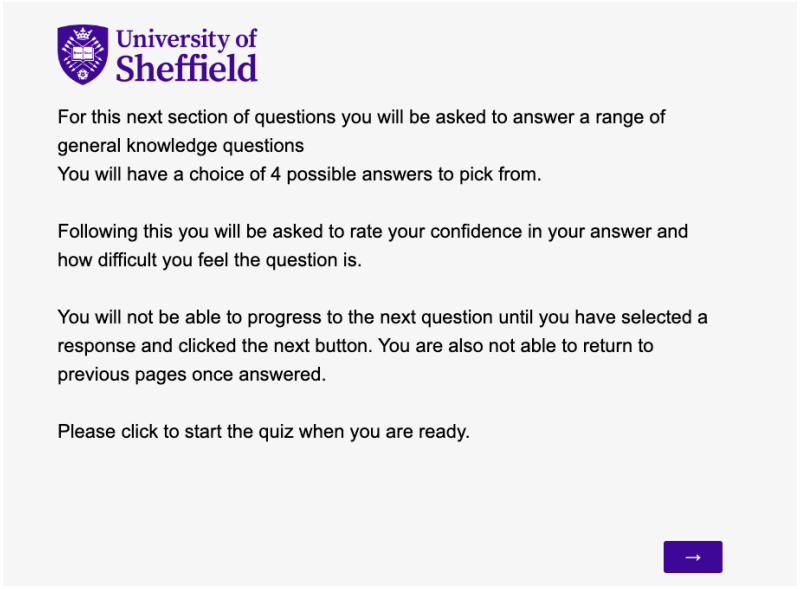
List of the 12 Threat Questions Used

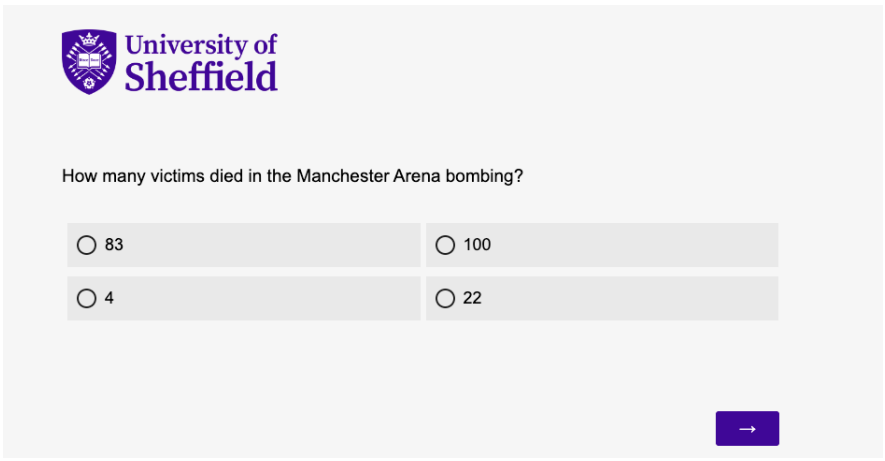
Question	Correct response	Other responses		
1. How many victims died in the Manchester Arena bombing?	**22**	83	100	4
2. What year did the first covid related lockdown start?	**2020**	2018	2021	2022
3. What year was the September 11 (9/11) attack?	**2001**	1997	2011	2003
4. Which country in Europe has highest number of prisoners per population?	**UK**	France	Spain	Italy
5. In which country did the Chernobyl nuclear disaster occur?	**Ukraine**	France	Belarus	Italy
6. What city in the UK has the highest reported knife crime?	**London**	Manchester	Birmingham	Glasgow
7. What is the term used to describe a type of malicious software designed to harm or exploit computer systems?	**Malware**	Spyware	Hardware	Airware
8. In the context of cybersecurity, what is the act of tricking someone into revealing sensitive information called?	**Phishing**	Spoofing	Hacking	DDoS
9. Which offence involves intentionally engaging in behaviour that causes another person to fear for their safety or the safety of others?	**Harassment**	Assault	Stalking	Fraud
10. How many people will develop some form of cancer during their lifetime?	**1 in 2**	1 in 10	1 in 8	1 in 5
11. Within a year (September 2021-2022) what % of adults were the victim of a crime?	**17%**	50%	22%	2%
12. Within two weeks in Feb 2023, how many in every 100 adults were finding it difficult to afford their rent or mortgage payments?	**32**	7	21	17

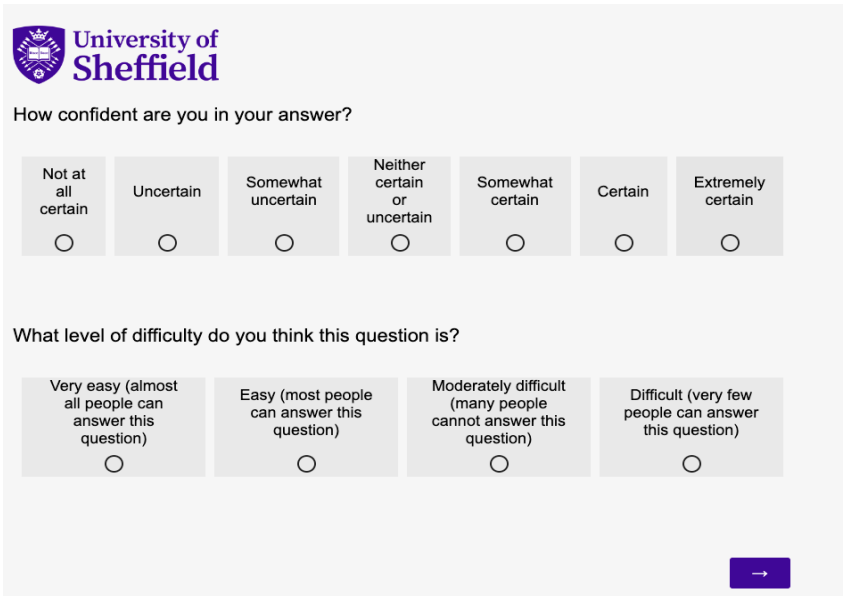
List of the 12 Neutral Questions Used

Question	Correct response	Other responses		
1. In 1768, Captain James Cook set out to explore which ocean?	**Pacific Ocean**	Atlantic Ocean	Indian Ocean	Arctic Ocean
2. Jeff Bezos is the founder of which billion-dollar company?	**Amazon**	Tesla	Apple	Subway
3. Lemurs are only native to one country, which one is it?	**Madagascar**	Peru	India	Cuba
4. In the UK, the abbreviation NHS stands for National what Service?	**Health**	Humanity	Honour	Household
5. What name is given to the revolving belt machinery in an airport that delivers checked luggage from the plane to baggage reclaim?	**Carousel**	Hanger	Terminal	Concourse
6. Which of these brands was chiefly associated with the manufacture of household locks?	**Chubb**	Phillips	Flymo	Ronseal
7. Construction of which of these famous landmarks was completed first?	**Big Ben Clock Tower**	Eiffel Tower	Royal Albert Hall	Empire State Building
8. What is the name of a female badger?	**A sow**	A boar	A vixen	A doe
9. How many moons does Mars have?	**2**	5	7	0
10. The Kodiak bear is native to which continent?	**North America**	Africa	Australia	Europe
11. What is the name of the 3000-mile-long mountain range that stretches through North America?	**Rockies**	Alpes	Andes	Himalayas
12. Which streaming service launched in the UK in 2020?	**Disney+**	Netflix	Amazon Prime Video	BBC i-player

Example Display of Instructions Page and an example of a Question Display


University of Sheffield
For this next section of questions you will be asked to answer a range of general knowledge questions
You will have a choice of 4 possible answers to pick from.
Following this you will be asked to rate your confidence in your answer and how difficult you feel the question is.
You will not be able to progress to the next question until you have selected a response and clicked the next button. You are also not able to return to previous pages once answered.
Please click to start the quiz when you are ready.
→


University of Sheffield
How many victims died in the Manchester Arena bombing?
83
100
4
22
→


University of Sheffield
How confident are you in your answer?
Not at all certain
Uncertain
Somewhat uncertain
Neither certain or uncertain
Somewhat certain
Certain
Extremely certain
What level of difficulty do you think this question is?
Very easy (almost all people can answer this question)
Easy (most people can answer this question)
Moderately difficult (many people cannot answer this question)
Difficult (very few people can answer this question)
→

Appendix O: Other Measures Collected as Part of the Larger Project

Multi-Thematic Delusional Beliefs Scale (MDBS; Martinez & Bentall, 2022)

The MDBS is a new scale developed to assess the presence of different delusional belief types (religious, grandiose, reference, control, and paranoid). Participants answered 40 questions measured on a five-point Likert scale (0= strongly disagree, 4= strongly agree).

Cohen's Social Network Index (CSNI; Cohen et al., 1997)

The CSNI measures network diversity, number of people in a network, and number of embedded networks people have. Participants answered questions on the CSNI.

AB Task

Participants first rated their how trustworthy and how threatened they felt by a person A and B (no other information given). Participants were then shown 10 pairs of neutral items and asked which they preferred. After each item they were then told either person A or B agreed with them. They were then asked to confirm who agreed with them. Person A will always agree with the participant 80% of the time. After all 10 pairs, participants were then shown a mystery item A and B and told person chose mystery Item A and person B chose mystery item B. The participant will then be asked to select an item of their choice. If the person has formed a coalition with person A they should pick mystery Item A. Participants then rated their trust in person A and B again. This task is repeated three further times using statements related to political beliefs, delusional beliefs and conspiracy theories.

Belief Exploration Task

Participants were shown a list of beliefs: 5 commonly held beliefs (the sun is a ball of gas),5 delusional beliefs taken from the Green et al. Paranoid Thought Scales (GPTS, 2008;

e.g., Jesus speaks to me); and 5 conspiracy theories (COVID-19 vaccinations are being used to shorten people's lives). For each belief, Participants were then asked:

In your opinion, how true do you think this belief is? (1-7 Likert scale).

1. How likely is it that others in your close circle will agree with you? (1-7 Likert scale).
2. How comfortable would you feel talking about this belief with others? (1-7 Likert scale).

At the end of the task participants were asked two further questions:

1. Rate your threat level of sharing these beliefs with others (1-7 Likert scale).
2. Rate your trust level of sharing these beliefs with others (1-7 Likert scale).

Multi-Thematic Delusional Beliefs Scale Questions

Please indicate the degree to which you agree or disagree with the following statements

	Religious delusional beliefs	*Strongly Disagree*				*Strongly Agree*
1	I am an important religious figure (e.g., a prophet or a messenger from God)	0	1	2	3	4
2	I have been chosen by a deity for a special role or mission	0	1	2	3	4
3	I am especially close to God or Gods	0	1	2	3	4
4	There's only one set of religious teachings in this world that is true and provides guidance about how to live	0	1	2	3	4
5	The fate of my immortal soul is determined by my deeds on Earth	0	1	2	3	4
6	Everything happens for a reason, as if my life has been planned by a higher power	0	1	2	3	4
7	It is helpful to seek guidance and comfort from a divine force	0	1	2	3	4
8	The natural world is so complex and perfect, it must have been designed by a superior being	0	1	2	3	4
	Grandiose delusional beliefs	*Strongly Disagree*				*Strongly Agree*
1	I have special abilities or powers that ordinary people do not have	0	1	2	3	4
2	I am connected with very important and powerful people	0	1	2	3	4
3	I am, or am destined to be, someone very important	0	1	2	3	4
4	I deserve to be seen as a great person	0	1	2	3	4
5	I tend to be the centre of attention because of my outstanding contributions	0	1	2	3	4
6	I have many special and unique ideas	0	1	2	3	4
7	Pretty much everything I do is great	0	1	2	3	4
8	Generally, I tend to perform better than other people	0	1	2	3	4
	Reference delusional beliefs	*Strongly Disagree*				*Strongly Agree*
1	Sometimes someone on the radio or television speaks directly to me.	0	1	2	3	4
2	I believe that articles in magazines, newspapers or on the internet are specially written for me.	0	1	2	3	4
3	I often think that people drop hints to me or say things with a double meaning	0	1	2	3	4
4	Public announcements made by politicians often seem to be specifically directed towards me	0	1	2	3	4
5	When I take a journey by bus or car traffic lights often change if I'm in a hurry	0	1	2	3	4
6	At times I mistakenly think that strangers are waving at me	0	1	2	3	4
7	Animals seem to take special notice of me when I go for a walk	0	1	2	3	4
8	Stories in books, films or TV shows often resemble my personal life story	0	1	2	3	4

	Control/passivity delusional beliefs	*Strongly Disagree*				*Strongly Agree*
1	There are times when I believe I am under the control of an external force or power, as if I were possessed	0	1	2	3	4
2	There are times when I think I act as a robot or zombie without a will of my own	0	1	2	3	4
3	At times, there is something inside of me that tries to take over my speech or actions	0	1	2	3	4
4	I have wondered whether some of my thoughts have been imposed on my mind by some force or external agent	0	1	2	3	4
5	Sometimes some of my thoughts seem so strange that I think they don't really belong to me.	0	1	2	3	4
6	Every so often I feel that my body isn't really mine	0	1	2	3	4
7	Sometimes I behave so differently, that I don't recognize myself	0	1	2	3	4
8	On occasions I act as if I were in automatic mode	0	1	2	3	4
	Paranoid delusional beliefs	*Strongly Disagree*				*Strongly Agree*
1	There are people who deliberately want to harm me.	0	1	2	3	4
2	Some people's intentions towards me are bad.	0	1	2	3	4
3	It is always important to be on guard against being deceived or harmed.	0	1	2	3	4
4	I suspect that people are watching me when I am out in public	0	1	2	3	4
5	People will almost certainly deceive me.	0	1	2	3	4
6	I expect I will be criticised or rejected in social situations	0	1	2	3	4
7	You should only trust yourself.	0	1	2	3	4
8	People sometimes talk about me behind my back.	0	1	2	3	4

Items based on the following scales:

***Religious delusional beliefs*:** Peter's delusion inventory (PDI-40; Peters et al., 1999)| Religiosity Scale (Alsuhibani, Shevlin, & Bentall, 2020)

***Grandiose delusional beliefs*:** Peter's delusion inventory (PDI-40; Peters et al., 1999) | Specific Psychotic Experiences Questionnaire (SPEQ; Ronald et al. 2014) |The Narcissistic Admiration and Rivalry Questionnaire-Short form (Leckelt et al., 2017)

***Self-reference delusional beliefs*:** Peter's delusion inventory (PDI-40; Peters et al., 1999) | Referential thinking scale (Lenzenweger et al., 1997) |Qualitative study report (Startup & Startup, 2005)

***Control/passivity delusional beliefs*:** Peter's delusion inventory (PDI-40; Peters et al., 1999) | The multidimensional inventory of dissociation (MID; Dell, 2006)

***Paranoid delusional beliefs*:** The persecution and deservedness scale (PaDS; Melo et al., 2009)

Cohen's Social Network Index Questions

Social Network Index

Instructions: This questionnaire is concerned with how many people you see or talk to on a regular basis including family, friends, workmates, neighbors, etc. Please read and answer each question carefully. Answer follow-up questions where appropriate.

1. Which of the following best describes your marital status?
   ___(1) currently married & living together, or living with someone in marital-like relationship
   ___(2) never married & never lived with someone in a marital-like relationship
   ___(3) separated
   ___(4) divorced or formerly lived with someone in a marital-like relationship
   ___(5) widowed

2. How many children do you have? (If you don't have any children, check '0' and skip to question 3.)

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

2a. How many of your children do you see or talk to on the phone at least once every 2 weeks?

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

3. Are either of your parents living? (If neither is living, check '0' and skip to question 4.)

____ (0) neither ____ (1) mother only ____ (2) father only ____ (3) both

3a. Do you see or talk on the phone to either of your parents at least once every 2 weeks?

____ (0) neither ____ (1) mother only ____ (2) father only ____ (3) both

4. Are either of your in-laws (or partner's parents) living? (If you have none, check the appropriate space and skip to question 5.)

____ (0) neither ____ (1) mother only ____ (2) father only ____ (3) both ____ (4) not applicable

4a. Do you see or talk on the phone to either of your partner's parents at least once every 2 weeks?

_____ (0) neither _____ (1) mother only _____ (2) father only ____ (3) both

5. How many other relatives (other than your spouse, parents & children) do you feel close to? (If '0', check that space and skip to question 6.)

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

5a. How many of these relatives do you see or talk to on the phone at least once every 2 weeks?

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

6. How many close friends do you have? (meaning people that you feel at ease with, can talk to about private matters, and can call on for help)

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

6a. How many of these friends do you see or talk to at least once every 2 weeks?

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

7. Do you belong to a church, temple, or other religious group? (If not, check 'no' and skip to question 8.)

_____ no _____ yes

7a. How many members of your church or religious group do you talk to at least once every 2 weeks? (This includes at group meetings and services.)

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

8. Do you attend any classes (school, university, technical training, or adult education) on a regular basis? (If not, check 'no' and skip to question 9.)

_____ no _____ yes

8a. How many fellow students or teachers do you talk to at least once every 2 weeks? (This includes at class meetings.)

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

9. Are you currently employed either full or part-time? (If not, check 'no' and skip to question 10.)

____ (0) no _____ (1) yes, self-employed _____ (2) yes, employed by others

9a. How many people do you supervise?

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

9b. How many people at work (other than those you supervise) do you talk to at least once every 2 weeks?

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

10. How many of your neighbors do you visit or talk to at least once every 2 weeks?

_____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

11. Are you currently involved in regular volunteer work? (If not, check 'no' and skip to question 12.)

_____ no _____ yes

11a. How many people involved in this volunteer work do you talk to about volunteering-related issues at least once every 2 weeks?

____0 ____1 ____2 ____3 ____4 ____5 ____6 ____7 or more

12. Do you belong to any groups in which you talk to one or more members of the group about group-related issues at least once every 2 weeks? Examples include social clubs, recreational groups, trade unions, commercial groups, professional organizations, groups concerned with children like the PTA or Boy Scouts, groups concerned with community service, etc. (If you don't belong to any such groups, check 'no' and skip the section below.)

_____ no _____ yes

Consider those groups in which you talk to a fellow group member at least once every 2 weeks. Please provide the following information for each such group: the name or type of group and the total number of members in that group that you talk to at least once every 2 weeks.

Group	Total number of group members that you talk to at least once every 2 weeks
1.	
2.	
3.	
4.	
5.	
6.	

This scale was used for the following journal article:

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M. Jr. (1997). Social ties and susceptibility to the common cold. Journal of the American Medical Association, 277, 1940-1944. Link to full-text (pdf)

Appendix P: Statistical Analyses Assumption Testing

Specific Tests of Assumptions Conducted

All data was assessed for outliers and normality. Outliers were identified (box plots) and reviewed for potential errors in data entry or measurement. Normality was assessed both visually (histograms, QQ plots) and statistically (Shapiro-Wilk's test $p > .05$). All violations to assumptions for specific tests are reported within the results section.

Descriptive tatistics

All outliers were deemed accurate and were included in the final data analyses.

One-Way Analysis of Variance (ANOVA)

Where assumptions of normality were violated, skewness within the groups was reviewed. In all cases, the decision was made to run one-way ANOVAs due to their robustness to deviations from normality, particularly with equal group sample sizes (Lix et al., 1996).

Homogeneity of variance was assessed using Levene's test of homogeneity of variances. Post hoc analysis (using pairwise comparisons) was performed to assess for significant differences between groups. ANOVA and Tukey post hoc tests are reported, except in instances where assumptions of homogeneity were violated or sample sizes were unequal. In these cases, a Welch ANOVA and Games-Howell post hoc tests are instead reported given their robustness/more conservative estimates where violations of homogeneity are present (Delacre et al., 2019).

Two-Way Mixed ANOVAS

Additional assumptions were checked. Homogeneity of variance-covariances matrices was assessed by Box's test of equality of covariance matrices ($p<.001$). Outliers were additionally assessed by studentized residual values (+-3). Assumptions of normality were also assessed based on residuals (QQ plots). Mauchly's test of sphericity for the assumption of sphericity was not violated in any case, given all within-subject factors only had two levels, as such interpretations were based on Sphericity Assumed. Main effects were explored with Bonferroni adjustment for confidence intervals.

Correlations.

Preliminary analyses were run assessing for a linear relationship between variables, outliers, and normal distribution of variables. All outliers were assessed as accurate, noted and included within analyses. Pearson's product-moment correlations were used for all correlations, including where data was non-normally distributed given its robustness to deviations from normality (Havlicek & Peterson, 1976).

Results of Violations and Meeting of Assumptions for Specific Analyses

Sample and Group Characteristics

Anxiety (GAD-7 Scores). No outliers were present in any group. Total anxiety score was normally distributed across all groups, however, assumptions of homogeneity of variances were violated ($p >.001$).

Depression (PHQ-9 Scores). Outliers were present in the GP group. Data was normally distributed in both the Del and MH groups ($p > .05$) but was not normally distributed in the GP group ($p<.001$). There was no homogeneity in variance ($p <.001$).

PDI Scores (Total and Yes/No). Outliers were present in the Del group for PDI total score however no outliers were found for PDI yes/no data. PDI total scores were not normally

distributed for any group and PDI yes/no scores were not normally distributed for Del and MH groups ($p < .05$). The assumption of homogeneity of variances was violated for both PDI total scores ($p<0.001$) and yes/no score ($p = .001$).

PANSS (Items P1, P5, P6). Outliers were found in data for items P1 (Del, GP), P5 (MH, GP), and P6 (Del). Data was non-normally distributed across all groups for all three item scores and the assumption of homogeneity of variances was violated for all three items ($p<.001$).

Whole Sample Analyses: Accuracy and Certainty.

Neutral: Correct and Incorrect. For all neutral questions, preliminary analyses showed the relationship to be linear, with one outlier present, however, accuracy scores violated assumptions of normality.

Neutral: Correct Only. For correct neutral questions only, preliminary analyses showed the relationship to be linear, with no outliers present, however, accuracy scores violated assumptions of normality.

Threat: Correct and Incorrect. Preliminary analyses showed the relationship to be linear, with no outliers present, however, accuracy scores violated assumptions of normality.

Threat: Correct Only. Preliminary analyses showed the relationship to be linear with outliers present, however, accuracy scores violated assumptions of normality.

Group Differences

Confidence Judgment Questionnaire. No outliers were present and assumptions of normality and homogeneity in variance were met.

Millionaire Game: Accuracy.

Threat vs Neutral. Three outliers were present in the GP group for threat questions and assumptions of normality were violated (in MH groups for threat questions) and homogeneity in variance was violated for neutral questions. There was one outlier for neutral accuracy data (studentized residual value of -3.19). There was homogeneity of variance-covariances matrices (p =.066).

Millionaire Game: Certainty.

Threat Questions: Incorrect vs Correct. No outliers were present. Assumptions of normality were violated (Del group for correct questions) and homogeneity in variance were violated for correct questions. There were no outliers (studentized residual values). There was homogeneity of variance-covariances matrices (p =.269).

Neutral Questions: Incorrect vs Correct. One outliner was present in the MH group for incorrect questions and for correct questions there were two outliers in MH group and three in GP group. Assumptions of normality were violated (MH group for incorrect questions) and assumptions of homogeneity in variance were met for both. There were no outliers (studentized residual values). There was homogeneity of variance-covariances matrices (p =.584).

Millionaire Game: Mean RT

Threat Questions: Incorrect vs Correct. Following transformation, three outliers were present in the MH group for incorrect RT data. Assumptions of normality were violated in the Del group for correct responses and in the MH group for incorrect questions only. Homogeneity in variance was violated for incorrect responses. There was one outlier for incorrect mean RTs (studentized residual value of 3.32). There was homogeneity of variance-covariances matrices (p =.43).

Neutral Questions: Incorrect vs Correct. Following transformation, one outlier was present in Del and GP groups for incorrect and one in Del group for correct. Assumptions of normality were violated in the Del group for correct and incorrect responses. Homogeneity in variance was violated for correct responses. There was one outlier for both incorrect and correct RTs (both studentized residual value of 3.07). There was homogeneity of variance-covariances matrices (p =.024).

Millionaire Game: Individual RT and Certainty Correlations

Threat vs Neutral. No outliers were present and assumptions of normality (in MH groups for both question types) and homogeneity in variance were violated (all groups). There were no outliers (studentized residual values). There was homogeneity of variance-covariances matrices (p =.025).

Appendix Q: SPSS Outputs

Sample and Group Characteristics

Age

Descriptives

Age

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	29.59	6.801	1.450	26.58	32.61	18	49
MH	22	33.18	10.866	2.317	28.36	38.00	21	65
GP	22	33.86	10.494	2.237	29.21	38.52	25	60
Total	66	32.21	9.604	1.182	29.85	34.57	18	65

ANOVA

Age

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	231.848	2	115.924	1.267	.289
Within Groups	5763.182	63	91.479		
Total	5995.030	65			

Gender

Gender * Which group, Delusion, MH, GP Crosstabulation

Count

		Which group, Delusion, MH, GP: Del	MH	GP	Total
Gender	Male	10_a	9_a	6_a	25
	Female	12_a	12_a	16_a	40
	3	0_a	1_a	0_a	1
Total		22	22	22	66

Each subscript letter denotes a subset of Which group, Delusion, MH, GP categories whose column proportions do not differ significantly from each other at the .05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	3.840[a]	4	.428
Likelihood Ratio	4.064	4	.397
Linear-by-Linear Association	1.368	1	.242
N of Valid Cases	66		

a. 3 cells (33.3%) have expected count less than 5. The minimum expected count is .33.

Ethnicity

Ethnicity * Which group, Delusion, MH, GP Crosstabulation

Count

		Which group, Delusion, MH, GP			
		Del	MH	GP	Total
Ethnicity	Asian or Asian British	2_a	0_a	0_a	2
	Black, Black British, Caribbean or African	10_a	1_b	0_b	11
	Mixed or multiple ethnic groups	0_a	2_a	2_a	4
	White	10_a	18_b	20_b	48
	ther ethnic group	0_a	1_a	0_a	1
Total		22	22	22	66

Each subscript letter denotes a subset of Which group, Delusion, MH, GP categories whose column proportions do not differ significantly from each other at the .05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	28.045[a]	8	<.001
Likelihood Ratio	31.069	8	<.001
Linear-by-Linear Association	16.227	1	<.001
N of Valid Cases	66		

a. 12 cells (80.0%) have expected count less than 5. The minimum expected count is .33.

Diagnosis

Crosstab

Count

		Which group, Delusion, MH, GP			
		Del	MH	GP	Total
Have a diagnosis yes/no	Yes	16$_a$	15$_a$	0$_b$	31
	No	6$_a$	7$_a$	22$_b$	35
Total		22	22	22	66

Each subscript letter denotes a subset of Which group, Delusion, MH, GP categories whose column proportions do not differ significantly from each other at the .05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	29.320[a]	2	<.001
Likelihood Ratio	37.949	2	<.001
Linear-by-Linear Association	23.005	1	<.001
N of Valid Cases	66		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.33.

Medication

Crosstab

Count

		Which group, Delusion, MH, GP			
		Del	MH	GP	Total
Are taking medication yes/no	Yes	10$_a$	11$_a$	0$_b$	21
	No	12$_a$	11$_a$	22$_b$	45
Total		22	22	22	66

Each subscript letter denotes a subset of Which group, Delusion, MH, GP categories whose column proportions do not differ significantly from each other at the .05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	15.505[a]	2	<.001
Likelihood Ratio	21.750	2	<.001
Linear-by-Linear Association	10.317	1	.001
N of Valid Cases	66		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.00.

Treatment/Therapy

Crosstab

Count

		Which group, Delusion, MH, GP			
		Del	MH	GP	Total
Have had therapy yes/no	Yes	15_a	14_a	2_b	31
	No	7_a	8_a	20_b	35
Total		22	22	22	66

Each subscript letter denotes a subset of Which group, Delusion, MH, GP categories whose column proportions do not differ significantly from each other at the .05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	19.100[a]	2	<.001
Likelihood Ratio	21.486	2	<.001
Linear-by-Linear Association	15.187	1	<.001
N of Valid Cases	66		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.33.

Religion

Crosstab

Count

		Which group, Delusion, MH, GP			
		Del	MH	GP	Total
Religion	Christian	11_a	6_a	6_a	23
	Buddhist	0_a	0_a	1_a	1
	Other	2_a	0_a	0_a	2
	Atheism	1_a	6_a	3_a	10
	Agnostic	0_a	3_a	2_a	5
	No Religion	7_a	7_a	10_a	24
	Prefer not to say	1_a	0_a	0_a	1
Total		22	22	22	66

Each subscript letter denotes a subset of Which group, Delusion, MH, GP categories whose column proportions do not differ significantly from each other at the .05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17.524[a]	12	.131
Likelihood Ratio	19.842	12	.070
Linear-by-Linear Association	1.790	1	.181
N of Valid Cases	66		

a. 15 cells (71.4%) have expected count less than 5. The minimum expected count is .33.

Anxiety

Descriptives

GAD_Total

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	9.64	4.635	.988	7.58	11.69	2	18
MH	22	9.95	3.199	.682	8.54	11.37	4	16
GP	22	3.09	2.114	.451	2.15	4.03	0	7
Total	66	7.56	4.674	.575	6.41	8.71	0	18

Robust Tests of Equality of Means

GAD_Total

	Statistic[a]	df1	df2	Sig.
Welch	43.389	2	38.455	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: GAD_Total

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	-.318	1.047	.950	-2.83	2.20
		GP	6.545*	1.047	<.001	4.03	9.06
	MH	Del	.318	1.047	.950	-2.20	2.83
		GP	6.864*	1.047	<.001	4.35	9.38
	GP	Del	-6.545*	1.047	<.001	-9.06	-4.03
		MH	-6.864*	1.047	<.001	-9.38	-4.35
Games-Howell	Del	MH	-.318	1.201	.962	-3.25	2.61
		GP	6.545*	1.086	<.001	3.87	9.23
	MH	Del	.318	1.201	.962	-2.61	3.25
		GP	6.864*	.818	<.001	4.87	8.86
	GP	Del	-6.545*	1.086	<.001	-9.23	-3.87
		MH	-6.864*	.818	<.001	-8.86	-4.87

*. The mean difference is significant at the 0.05 level.

Depression

Descriptives

PHQ_Total

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	10.45	6.045	1.289	7.77	13.13	1	24
MH	22	11.64	6.298	1.343	8.84	14.43	3	25
GP	22	2.36	2.401	.512	1.30	3.43	0	9
Total	66	8.15	6.613	.814	6.53	9.78	0	25

Robust Tests of Equality of Means

PHQ_Total

	Statistic[a]	df1	df2	Sig.
Welch	32.831	2	34.654	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: PHQ_Total

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	-1.182	1.576	.735	-4.97	2.60
		GP	8.091*	1.576	<.001	4.31	11.87
	MH	Del	1.182	1.576	.735	-2.60	4.97
		GP	9.273*	1.576	<.001	5.49	13.06
	GP	Del	-8.091*	1.576	<.001	-11.87	-4.31
		MH	-9.273*	1.576	<.001	-13.06	-5.49
Games-Howell	Del	MH	-1.182	1.861	.802	-5.70	3.34
		GP	8.091*	1.387	<.001	4.66	11.53
	MH	Del	1.182	1.861	.802	-3.34	5.70
		GP	9.273*	1.437	<.001	5.71	12.84
	GP	Del	-8.091*	1.387	<.001	-11.53	-4.66
		MH	-9.273*	1.437	<.001	-12.84	-5.71

*. The mean difference is significant at the 0.05 level.

PDI

Descriptives

PDI_Total

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	110.82	56.914	12.134	85.58	136.05	43	278
MH	22	30.77	22.069	4.705	20.99	40.56	0	70
GP	22	16.64	13.131	2.800	10.81	22.46	0	38
Total	66	52.74	54.825	6.748	39.26	66.22	0	278

Robust Tests of Equality of Means

PDI_Total

	Statistic[a]	df1	df2	Sig.
Welch	29.534	2	35.676	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: PDI_Total

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	80.045*	10.869	<.001	53.96	106.14
		GP	94.182*	10.869	<.001	68.09	120.27
	MH	Del	-80.045*	10.869	<.001	-106.14	-53.96
		GP	14.136	10.869	.400	-11.95	40.23
	GP	Del	-94.182*	10.869	<.001	-120.27	-68.09
		MH	-14.136	10.869	.400	-40.23	11.95
Games-Howell	Del	MH	80.045*	13.014	<.001	47.79	112.30
		GP	94.182*	12.453	<.001	63.02	125.35
	MH	Del	-80.045*	13.014	<.001	-112.30	-47.79
		GP	14.136*	5.475	.037	.72	27.55
	GP	Del	-94.182*	12.453	<.001	-125.35	-63.02
		MH	-14.136*	5.475	.037	-27.55	-.72

*. The mean difference is significant at the 0.05 level.

Descriptives

PDI_Yes

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	10.95	3.836	.818	9.25	12.66	8	20
MH	22	3.55	1.945	.415	2.68	4.41	0	6
GP	22	2.86	2.007	.428	1.97	3.75	0	6
Total	66	5.79	4.573	.563	4.66	6.91	0	20

Robust Tests of Equality of Means

PDI_Yes

	Statistic[a]	df1	df2	Sig.
Welch	39.534	2	39.890	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: PDI_Yes

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	7.409*	.826	<.001	5.43	9.39
		GP	8.091*	.826	<.001	6.11	10.07
	MH	Del	-7.409*	.826	<.001	-9.39	-5.43
		GP	.682	.826	.689	-1.30	2.66
	GP	Del	-8.091*	.826	<.001	-10.07	-6.11
		MH	-.682	.826	.689	-2.66	1.30
Games-Howell	Del	MH	7.409*	.917	<.001	5.15	9.67
		GP	8.091*	.923	<.001	5.82	10.36
	MH	Del	-7.409*	.917	<.001	-9.67	-5.15
		GP	.682	.596	.493	-.77	2.13
	GP	Del	-8.091*	.923	<.001	-10.36	-5.82
		MH	-.682	.596	.493	-2.13	.77

*. The mean difference is significant at the 0.05 level.

PANSS P1

Descriptives

PANSS_P1_Del

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	3.68	.945	.202	3.26	4.10	3	6
MH	22	1.27	.456	.097	1.07	1.47	1	2
GP	22	1.14	.351	.075	.98	1.29	1	2
Total	66	2.03	1.336	.164	1.70	2.36	1	6

Robust Tests of Equality of Means

PANSS_P1_Del

	Statistic[a]	df1	df2	Sig.
Welch	69.880	2	38.426	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: PANSS_P1_Del

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	2.409*	.193	<.001	1.95	2.87
		GP	2.545*	.193	<.001	2.08	3.01
	MH	Del	-2.409*	.193	<.001	-2.87	-1.95
		GP	.136	.193	.760	-.33	.60
	GP	Del	-2.545*	.193	<.001	-3.01	-2.08
		MH	-.136	.193	.760	-.60	.33
Games-Howell	Del	MH	2.409*	.224	<.001	1.86	2.96
		GP	2.545*	.215	<.001	2.01	3.08
	MH	Del	-2.409*	.224	<.001	-2.96	-1.86
		GP	.136	.123	.513	-.16	.44
	GP	Del	-2.545*	.215	<.001	-3.08	-2.01
		MH	-.136	.123	.513	-.44	.16

*. The mean difference is significant at the 0.05 level.

PANSS P5

Descriptives

PANSS_P5_Gran

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	2.59	1.501	.320	1.93	3.26	1	5
MH	22	1.18	.395	.084	1.01	1.36	1	2
GP	22	1.05	.213	.045	.95	1.14	1	2
Total	66	1.61	1.135	.140	1.33	1.89	1	5

Robust Tests of Equality of Means

PANSS_P5_Gran

	Statistic[a]	df1	df2	Sig.
Welch	11.814	2	34.340	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: PANSS_P5_Gran

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	1.409*	.273	<.001	.75	2.06
		GP	1.545*	.273	<.001	.89	2.20
	MH	Del	-1.409*	.273	<.001	-2.06	-.75
		GP	.136	.273	.872	-.52	.79
	GP	Del	-1.545*	.273	<.001	-2.20	-.89
		MH	-.136	.273	.872	-.79	.52
Games-Howell	Del	MH	1.409*	.331	<.001	.58	2.24
		GP	1.545*	.323	<.001	.73	2.36
	MH	Del	-1.409*	.331	<.001	-2.24	-.58
		GP	.136	.096	.340	-.10	.37
	GP	Del	-1.545*	.323	<.001	-2.36	-.73
		MH	-.136	.096	.340	-.37	.10

*. The mean difference is significant at the 0.05 level.

PANSS P6

Descriptives

PANSS_P6_Per

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Del	22	3.36	1.255	.268	2.81	3.92	1	5
MH	22	1.50	.512	.109	1.27	1.73	1	2
GP	22	1.32	.477	.102	1.11	1.53	1	2
Total	66	2.06	1.239	.152	1.76	2.37	1	5

Robust Tests of Equality of Means

PANSS_P6_Per

	Statistic[a]	df1	df2	Sig.
Welch	25.231	2	38.938	<.001

a. Asymptotically F distributed.

Multiple Comparisons

Dependent Variable: PANSS_P6_Per

	(I) Group_Number	(J) Group_Number	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Del	MH	1.864*	.250	<.001	1.26	2.46
		GP	2.045*	.250	<.001	1.45	2.65
	MH	Del	-1.864*	.250	<.001	-2.46	-1.26
		GP	.182	.250	.749	-.42	.78
	GP	Del	-2.045*	.250	<.001	-2.65	-1.45
		MH	-.182	.250	.749	-.78	.42
Games-Howell	Del	MH	1.864*	.289	<.001	1.15	2.58
		GP	2.045*	.286	<.001	1.34	2.76
	MH	Del	-1.864*	.289	<.001	-2.58	-1.15
		GP	.182	.149	.449	-.18	.54
	GP	Del	-2.045*	.286	<.001	-2.76	-1.34
		MH	-.182	.149	.449	-.54	.18

*. The mean difference is significant at the 0.05 level.

Whole Sample

Accuracy and Certainty

Neutral: Correct and Incorrect.

Correlations

		M_CERT_N	Total number correct answers on neutral questions (12)
M_CERT_N	Pearson Correlation	1	.549**
	Sig. (2-tailed)		<.001
	N	66	66
Total number correct answers on neutral questions (12)	Pearson Correlation	.549**	1
	Sig. (2-tailed)	<.001	
	N	66	66

**. Correlation is significant at the 0.01 level (2-tailed).

Neutral: Correct Only.

Correlations

		Total number correct answers on neutral questions (12)	CO_M_CERT_N
Total number correct answers on neutral questions (12)	Pearson Correlation	1	.445**
	Sig. (2-tailed)		<.001
	N	66	66
CO_M_CERT_N	Pearson Correlation	.445**	1
	Sig. (2-tailed)	<.001	
	N	66	66

**. Correlation is significant at the 0.01 level (2-tailed).

Threat: Correct and Incorrect

Correlations

		Total number correct answers on threat questions (12)	M_CERT_T
Total number correct answers on threat questions (12)	Pearson Correlation	1	.281*
	Sig. (2-tailed)		.022
	N	66	66
M_CERT_T	Pearson Correlation	.281*	1
	Sig. (2-tailed)	.022	
	N	66	66

*. Correlation is significant at the 0.05 level (2-tailed).

Threat: Correct Only.

Correlations

		Total number correct answers on threat questions (12)	CO_M_CERT_T
Total number correct answers on threat questions (12)	Pearson Correlation	1	.143
	Sig. (2-tailed)		.251
	N	66	66
CO_M_CERT_T	Pearson Correlation	.143	1
	Sig. (2-tailed)	.251	
	N	66	66

Group Differences: CJQ

Descriptive Statistics

Dependent Variable: Total certainty score confidence judgment questionnaire

Which group, Delusion, MH, GP	Mean	Std. Deviation	N
Del	37.00	6.810	22
MH	37.23	4.810	22
GP	37.91	4.587	22
Total	37.38	5.423	66

Tests of Between-Subjects Effects

Dependent Variable: Total certainty score confidence judgment questionnaire

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	9.848[a]	2	4.924	.163	.850	.005
Intercept	92213.470	1	92213.470	3054.900	<.001	.980
Group_Number	9.848	2	4.924	.163	.850	.005
Error	1901.682	63	30.185			
Total	94125.000	66				
Corrected Total	1911.530	65				

a. R Squared = .005 (Adjusted R Squared = -.026)

Group Differences: Millionaire Game

Accuracy. Threat vs Neutral

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Threat_Neutral	Sphericity Assumed	5.121	1	5.121	2.773	.101	.042
	Greenhouse-Geisser	5.121	1.000	5.121	2.773	.101	.042
	Huynh-Feldt	5.121	1.000	5.121	2.773	.101	.042
	Lower-bound	5.121	1.000	5.121	2.773	.101	.042
Threat_Neutral * Group_Number	Sphericity Assumed	5.515	2	2.758	1.493	.233	.045
	Greenhouse-Geisser	5.515	2.000	2.758	1.493	.233	.045
	Huynh-Feldt	5.515	2.000	2.758	1.493	.233	.045
	Lower-bound	5.515	2.000	2.758	1.493	.233	.045
Error(Threat_Neutral)	Sphericity Assumed	116.364	63	1.847			
	Greenhouse-Geisser	116.364	63.000	1.847			
	Huynh-Feldt	116.364	63.000	1.847			
	Lower-bound	116.364	63.000	1.847			

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	9742.091	1	9742.091	1848.650	<.001	.967
Group_Number	68.909	2	34.455	6.538	.003	.172
Error	332.000	63	5.270			

Pairwise Comparisons

Measure: MEASURE_1

(I) Which group, Delusion, MH, GP	(J) Which group, Delusion, MH, GP	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
Del	MH	-1.318*	.489	.027	-2.522	-.114
	GP	-1.682*	.489	.003	-2.886	-.478
MH	Del	1.318*	.489	.027	.114	2.522
	GP	-.364	.489	1.000	-1.567	.840
GP	Del	1.682*	.489	.003	.478	2.886
	MH	.364	.489	1.000	-.840	1.567

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Mean Certainty

Threat: Correct vs Incorrect.

Tests of Within-Subjects Effects

Measure: Cert

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
correct_incorrect_threat	Sphericity Assumed	78.436	1	78.436	103.258	<.001	.665
	Greenhouse-Geisser	78.436	1.000	78.436	103.258	<.001	.665
	Huynh-Feldt	78.436	1.000	78.436	103.258	<.001	.665
	Lower-bound	78.436	1.000	78.436	103.258	<.001	.665
correct_incorrect_threat * Group_Number	Sphericity Assumed	3.678	2	1.839	2.421	.099	.085
	Greenhouse-Geisser	3.678	2.000	1.839	2.421	.099	.085
	Huynh-Feldt	3.678	2.000	1.839	2.421	.099	.085
	Lower-bound	3.678	2.000	1.839	2.421	.099	.085
Error(correct_incorrect_threat)	Sphericity Assumed	39.500	52	.760			
	Greenhouse-Geisser	39.500	52.000	.760			
	Huynh-Feldt	39.500	52.000	.760			
	Lower-bound	39.500	52.000	.760			

Tests of Between-Subjects Effects

Measure: Cert

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1168.710	1	1168.710	783.396	<.001	.938
Group_Number	6.146	2	3.073	2.060	.138	.073
Error	77.576	52	1.492			

Neutral: Correct vs Incorrect

Tests of Within-Subjects Effects

Measure: Cert

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
correct_incorrect_neutral	Sphericity Assumed	111.599	1	111.599	174.290	<.001	.784
	Greenhouse-Geisser	111.599	1.000	111.599	174.290	<.001	.784
	Huynh-Feldt	111.599	1.000	111.599	174.290	<.001	.784
	Lower-bound	111.599	1.000	111.599	174.290	<.001	.784
correct_incorrect_neutral * Group_Number	Sphericity Assumed	2.667	2	1.333	2.082	.136	.080
	Greenhouse-Geisser	2.667	2.000	1.333	2.082	.136	.080
	Huynh-Feldt	2.667	2.000	1.333	2.082	.136	.080
	Lower-bound	2.667	2.000	1.333	2.082	.136	.080
Error(correct_incorrect_neutral)	Sphericity Assumed	30.735	48	.640			
	Greenhouse-Geisser	30.735	48.000	.640			
	Huynh-Feldt	30.735	48.000	.640			
	Lower-bound	30.735	48.000	.640			

Tests of Between-Subjects Effects

Measure: Cert

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	773.712	1	773.712	466.791	<.001	.907
Group_Number	.826	2	.413	.249	.781	.010
Error	79.561	48	1.658			

Pairwise Comparisons

Measure: Cert

(I) correct_incorrect_neutral	(J) correct_incorrect_neutral	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
1	2	2.102*	.159	<.001	1.782	2.422
2	1	-2.102*	.159	<.001	-2.422	-1.782

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Mean RT

Threat: Correct vs Incorrect

Tests of Within-Subjects Effects

Measure: Cert

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
correct_incorrect_neutral	Sphericity Assumed	.005	1	.005	2.737	.104	.050
	Greenhouse-Geisser	.005	1.000	.005	2.737	.104	.050
	Huynh-Feldt	.005	1.000	.005	2.737	.104	.050
	Lower-bound	.005	1.000	.005	2.737	.104	.050
correct_incorrect_neutral * Group_Number	Sphericity Assumed	.006	2	.003	1.779	.179	.064
	Greenhouse-Geisser	.006	2.000	.003	1.779	.179	.064
	Huynh-Feldt	.006	2.000	.003	1.779	.179	.064
	Lower-bound	.006	2.000	.003	1.779	.179	.064
Error(correct_incorrect_neutral)	Sphericity Assumed	.095	52	.002			
	Greenhouse-Geisser	.095	52.000	.002			
	Huynh-Feldt	.095	52.000	.002			
	Lower-bound	.095	52.000	.002			

Tests of Between-Subjects Effects

Measure: Cert

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1.755	1	1.755	238.187	<.001	.821
Group_Number	.128	2	.064	8.717	<.001	.251
Error	.383	52	.007			

Non-Transformed Values.

Descriptive Statistics

	Which group, Delusion, MH, GP	Mean	Std. Deviation	N
CO_M_RT_T	Del	19.84866	13.791362	20
	MH	9.36006	9.746818	17
	GP	8.09749	5.187249	18
	Total	12.76089	11.534220	55
IO_Mean_T_RT	Del	18.84	16.629	20
	MH	11.83	10.402	17
	GP	8.69	3.636	18
	Total	13.35	12.355	55

Pairwise Comparisons

Measure: Cert

(I) Which group, Delusion, MH, GP	(J) Which group, Delusion, MH, GP	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
Del	MH	8.751*	3.252	.029	.705	16.796
	GP	10.952*	3.203	.004	3.028	18.876
MH	Del	-8.751*	3.252	.029	-16.796	-.705
	GP	2.202	3.334	1.000	-6.047	10.450
GP	Del	-10.952*	3.203	.004	-18.876	-3.028
	MH	-2.202	3.334	1.000	-10.450	6.047

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Neutral: Correct vs Incorrect.

Tests of Within-Subjects Effects

Measure: RT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
correct_incorrect_neutral	Sphericity Assumed	.001	1	.001	.499	.483	.010
	Greenhouse-Geisser	.001	1.000	.001	.499	.483	.010
	Huynh-Feldt	.001	1.000	.001	.499	.483	.010
	Lower-bound	.001	1.000	.001	.499	.483	.010
correct_incorrect_neutral * Group_Number	Sphericity Assumed	.001	2	.001	.284	.754	.012
	Greenhouse-Geisser	.001	2.000	.001	.284	.754	.012
	Huynh-Feldt	.001	2.000	.001	.284	.754	.012
	Lower-bound	.001	2.000	.001	.284	.754	.012
Error(correct_incorrect_neutral)	Sphericity Assumed	.122	48	.003			
	Greenhouse-Geisser	.122	48.000	.003			
	Huynh-Feldt	.122	48.000	.003			
	Lower-bound	.122	48.000	.003			

Tests of Between-Subjects Effects

Measure: RT

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1.804	1	1.804	303.144	<.001	.863
Group_Number	.051	2	.026	4.305	.019	.152
Error	.286	48	.006			

Non-Transformed Values.

Descriptive Statistics

	Which group, Delusion, MH, GP	Mean	Std. Deviation	N
CO_M_RT_N	Del	17.06077	13.139584	19
	MH	8.79686	6.232341	17
	GP	7.14767	2.156851	15
	Total	11.39052	9.787285	51
IO_Mean_N_RT	Del	12.7595	7.22981	19
	MH	8.3090	4.03446	17
	GP	7.8764	2.22304	15
	Total	9.8398	5.53194	51

Pairwise Comparisons

Measure: RT

(I) Which group, Delusion, MH, GP	(J) Which group, Delusion, MH, GP	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
Del	MH	6.357*	1.925	.005	1.582	11.133
	GP	7.398*	1.992	.002	2.458	12.339
MH	Del	-6.357*	1.925	.005	-11.133	-1.582
	GP	1.041	2.043	1.000	-4.026	6.108
GP	Del	-7.398*	1.992	.002	-12.339	-2.458
	MH	-1.041	2.043	1.000	-6.108	4.026

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Individual RT and Certainty Correlations Correct Only: Threat Vs Neutral

Descriptive Statistics

	Which group, Delusion, MH, GP	Mean	Std. Deviation	N
Correct only T_RTxCer - Correlation within-p RT and certainty rating threat questions	Del	-.085854	.6062475	22
	MH	-.409421	.4252633	22
	GP	-.385748	.3397703	22
	Total	-.293674	.4863018	66
Correct only N_RTxCer- Correlation within-p RT and certainty rating neutral questions	Del	.014853	.5094584	22
	MH	-.396950	.3048319	22
	GP	-.296753	.2885632	22
	Total	-.226284	.4147250	66

Tests of Within-Subjects Effects

Measure: CO_RTxCert

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
threat_neutral	Sphericity Assumed	.150	1	.150	1.665	.202	.026
	Greenhouse-Geisser	.150	1.000	.150	1.665	.202	.026
	Huynh-Feldt	.150	1.000	.150	1.665	.202	.026
	Lower-bound	.150	1.000	.150	1.665	.202	.026
threat_neutral * Group_Number	Sphericity Assumed	.051	2	.025	.281	.756	.009
	Greenhouse-Geisser	.051	2.000	.025	.281	.756	.009
	Huynh-Feldt	.051	2.000	.025	.281	.756	.009
	Lower-bound	.051	2.000	.025	.281	.756	.009
Error(threat_neutral)	Sphericity Assumed	5.672	63	.090			
	Greenhouse-Geisser	5.672	63.000	.090			
	Huynh-Feldt	5.672	63.000	.090			
	Lower-bound	5.672	63.000	.090			

Tests of Between-Subjects Effects

Measure: CO_RTxCert

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	8.922	1	8.922	32.267	<.001	.339
Group_Number	3.410	2	1.705	6.167	.004	.164
Error	17.419	63	.276			

Pairwise Comparisons

Measure: CO_RTxCert

(I) Which group, Delusion, MH, GP	(J) Which group, Delusion, MH, GP	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
Del	MH	.368*	.112	.005	.092	.643
	GP	.306*	.112	.025	.030	.581
MH	Del	-.368*	.112	.005	-.643	-.092
	GP	-.062	.112	1.000	-.338	.214
GP	Del	-.306*	.112	.025	-.581	-.030
	MH	.062	.112	1.000	-.214	.338

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Sensitivity Analyses: Both Correct and Incorrect Individual RT and Certainty Correlations: Threat Vs Neutral

Descriptive Statistics

	Which group, Delusion, MH, GP	Mean	Std. Deviation	N
Correlation within-p RT and certainty threat questions	Del	-.060564	.4459956	22
	MH	-.383121	.3901265	22
	GP	-.376677	.2861753	22
	Total	-.273454	.4036229	66
Correlation within-p RT and certainty neutral questions	Del	.078942	.4789771	22
	MH	-.376942	.3299644	22
	GP	-.282900	.2680991	22
	Total	-.193633	.4144082	66

Tests of Within-Subjects Effects

Measure: RTxcert

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
threat_neutral	Sphericity Assumed	.210	1	.210	3.014	.087	.046
	Greenhouse-Geisser	.210	1.000	.210	3.014	.087	.046
	Huynh-Feldt	.210	1.000	.210	3.014	.087	.046
	Lower-bound	.210	1.000	.210	3.014	.087	.046
threat_neutral * Group_Number	Sphericity Assumed	.101	2	.050	.724	.489	.022
	Greenhouse-Geisser	.101	2.000	.050	.724	.489	.022
	Huynh-Feldt	.101	2.000	.050	.724	.489	.022
	Lower-bound	.101	2.000	.050	.724	.489	.022
Error(threat_neutral)	Sphericity Assumed	4.395	63	.070			
	Greenhouse-Geisser	4.395	63.000	.070			
	Huynh-Feldt	4.395	63.000	.070			
	Lower-bound	4.395	63.000	.070			

Tests of Between-Subjects Effects

Measure: RTxcert

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	7.200	1	7.200	34.073	<.001	.351
Group_Number	3.944	2	1.972	9.333	<.001	.229
Error	13.312	63	.211			

Pairwise Comparisons

Measure: RTxcert

(I) Which group, Delusion, MH, GP	(J) Which group, Delusion, MH, GP	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	Upper Bound
Del	MH	.389*	.098	<.001	.148	.630
	GP	.339*	.098	.003	.098	.580
MH	Del	-.389*	.098	<.001	-.630	-.148
	GP	-.050	.098	1.000	-.291	.191
GP	Del	-.339*	.098	.003	-.580	-.098
	MH	.050	.098	1.000	-.191	.291

Based on estimated marginal means*. The mean difference is significant at the .05 level. b. Adjustment for multiple comparisons: Bonferroni.