Belief Flexibility and Trustworthiness in Adults Experiencing Paranoia

Bejal Fatania

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of

Clinical Psychology

The University of Sheffield

Faculty of Science

Clinical Psychology Unit, Department of Psychology

May 2022

Declaration

I declare that this work has not been submitted for any other degree or to any other institution.

Word Counts

Section One: Literature Review

Excluding references and tables: 6859

Including references and tables: 9426

Section Two: Empirical Study

Excluding references and tables: 8000

Including references and tables: 11,579

Total

Excluding references and tables: 14,859

Including references and tables: 21,005

Lay Summary

Previous research has found associations between belief flexibility and delusional severity in clinical and non-clinical samples, but little is known about paranoia and belief flexibility. Additionally, studies have found people with paranoia make lower trustworthiness judgments compared to healthy controls, however, the research lacks relevance to day-to-day life. Section one of this thesis aimed to review the current literature and develop an understanding of the relationship between belief flexibility and paranoia. Moreover, the review hoped to gain insight into whether there was a difference in belief flexibility in individuals experiencing low and high paranoia. A systematic literature search identified eight studies to be included in the current review. Of the eight studies, six studies examined non-clinical samples whilst two studies assessed clinical samples with non-clinical samples recruited as controls. Despite efforts to identify studies investigating differences in belief flexibility between low and high paranoia groups, only three studies were found. The review yielded mixed findings on the relationship between belief flexibility and paranoia in non-clinical and clinical samples, with some evidence suggesting that reduced belief flexibility predicted paranoia. However, of the three studies examining the difference in belief flexibility between the two groups, only one study reported on the association, which was found to be non-significant. The current review was the first to examine the relationship between belief flexibility and paranoia. However, due to the methodological limitations of the included studies, the findings from this review should be interpreted with caution. Further studies are required to aid the understanding of belief flexibility in people experiencing paranoia.

The second section of the thesis focused on the empirical study. The current study aimed to identify whether there was a difference in the rate a person experiencing low or high paranoia adjusts their trust judgements after receiving new information. Participants completed an online scenarios-based task measuring trust judgements in trustworthy and untrustworthy

conditions. Questionnaires related to paranoia, attachment, self-esteem, analytical thinking, and belief updating were also completed. The study examined changes in trustworthiness judgements when the characters in the task behaved consistently, for example, always trustworthy or untrustworthy, and following a trust violation when the characters suddenly behaved in an opposite fashion to their previous behaviour. Trustworthiness judgements in both conditions between the low and high paranoia group were non-significant when the characters behaved consistently and following a trust violation. However, participants changed their trust judgements in response to additional information in the way that was expected, suggesting the task in the current study is sensitive to measures of trust judgements. In both conditions, a significant interaction between gender and time, with greater trustworthiness rating for females across time, and a significant interaction between age and group, with greater paranoia severity demonstrated in younger people, was found. However, the association between high paranoia and attachment insecurity and self-esteem was non-significant. The strengths and limitations of the study, implications of the findings, and recommendations for future research are addressed in the study.

Acknowledgements

I would like to thank my supervisor Professor Richard Bentall for your ideas and support throughout the project. I am extremely grateful to Anton Martinez for supporting me throughout the process of this project and for our coffee mornings.

I would also like to thank all the participants who took the time to complete the study, this project could not have been completed without you all.

Thank you to all my friends and fellow trainees who encouraged me, motivated me, and kept me going when things were hard.

I would like to express my love and thanks to my grandparents who sadly passed at the start of training but always encouraged me to follow my dreams and to never give up.

Thank you to my partner, Jordan, for your love and continuous support. I appreciate your patience and encouragement, and that you always found a way to make me smile when the course became hard.

A huge thank you to my sister, Pooja, who has provided endless laughter, hope, and motivation over the last three years. Thank you for the emotional support and for putting work and life into perspective.

Finally, to my parents, Kanu and Champa. Your love and support along this journey are unmatched. I am incredibly grateful for everything you have done for me to help me get to where I am today. Thank you for believing in me and encouraging me to be the best version of myself. I dedicate this thesis to you both.

Table of Contents

Section One: Literature Review

Belief Flexibility in Adults Experiencing Paranoia: A Systematic Review

Abstract

Objectives

Belief flexibility has been studied extensively across clinical and non-clinical delusion, however, little is known about the specific delusional symptoms, for example, paranoia and the relationship with belief flexibility. The current systematic review aimed to understand the relationship between belief flexibility and paranoia in adults and to identify whether a difference in belief flexibility occurred within low and high paranoia groups.

Methods

A systematic search of three databases, Scopus, PsycINFO, and MEDLINE, was conducted in February 2022 using pre-defined search terms. Articles examining the relationship between belief flexibility and paranoia in adults were included in the review. A narrative synthesis of the results was completed, and the included studies were appraised for quality.

Results

Eight studies were included in the review. Global rating for the individual studies was predominantly 'weak' with two studies rated as 'moderate'. The association between belief flexibility and paranoia across non-clinical and clinical samples was inconsistent, however, one study did demonstrate that reduced belief flexibility predicts paranoia. Only one study reported the difference in belief inflexibility and paranoia, nevertheless, the association was non-significant.

Conclusions

The current review found some evidence for a relationship between belief inflexibility and paranoia in adults but there were limited studies examining differences in belief flexibility in low and high paranoia groups. Overall, methodological shortcomings of the individual

studies and inconsistent findings make it difficult to draw meaningful conclusions. Strengths and limitations of the review, implications of the findings, and future research direction are discussed in the review.

Keywords: paranoia, persecutory, belief flexibility, belief updating, belief formation.

Practitioner points:

- There is some evidence indicating an association between reduced belief flexibility and paranoia, however, high-quality research is required to conclude with confidence.

- There are a limited number of studies looking at differences in belief flexibility for low and high paranoia groups

Introduction

Belief flexibility can be defined as the degree to which a person demonstrates flexibility in their strongly held belief (Garety et al., 2020). The construct of belief flexibility emerged following early research focusing on cognitive biases, namely, jumping to conclusions (JTC) bias. It became apparent that whilst JTC assessed a data-gathering bias, the construct of belief flexibility was more sophisticated and involved a meta-cognitive reasoning capacity (Garety et al., 2005). The process of belief flexibility requires an individual to uncouple from a strongly held belief by considering the likelihood of being mistaken, reflecting on their held belief when new evidence and/or information is presented (evidence integration); and generating and/or considering alternative explanations (Fischoff & Beythmarom, 1983, Hemsley & Garety, 1986, Ward & Garety, 2019).

Measures of Belief Flexibility

Belief flexibility, or the inverse belief inflexibility, is measured using a variety of methods including clinical interviews or delusion-neutral tasks. The various measures of belief flexibility will be discussed in this section to gain an understanding of the available tools. However, the literature review will only focus on the iterations of BADE as it is the only experimental measure to assess belief flexibility and the only one for which there was a sufficient number of studies specifically focusing on paranoia. Regarding the clinical interview, assessment for belief flexibility is conducted using the Maudsley Assessment of Delusions Schedule (MADS; Wessely et al., 1993) and the Explanations of Experiences (EoE) interview assessing Alternative Explanations (AE; Freeman et al., 2004).

The MADS is a 53-item interview assessing eight dimensions of delusions and an individual's reasoning about their experiences and strongly held beliefs. The dimensions of delusions include: (1) strength of conviction, (2) belief maintenance, (3) affect related to

belief, (4) action on beliefs, (5) idiosyncrasy of belief, (6) preoccupation with belief, (7) systematization of belief, and (8) insight.

Two items in the belief maintenance section are used to assess elements of belief flexibility, the possibility of being mistaken (PM), for example, "When you think about it now is it at all possible that you are mistaken about X?" and the reaction to a hypothetical contradiction (RTHC), for example, "How would you react in a hypothetical situation if new evidence were to be generated which contradicts the delusion". Using this approach belief flexibility is operationalised as the delusional belief being absent or present, which is rated as 'yes' or 'no', respectively. The MADS has very good inter-rater reliability with excellent kappa values for PM and RTHC (.91 and .90, respectively; Wessely et al., 1993).

The EoE interview (Freeman et al., 2004) assesses an individual's ability to consider an alternate explanation for their delusions. Individuals are typically asked, "Can you think of any other explanations for the experiences that you have described?" and "Are there any other reasons - other than [the delusional belief] - that could account for these experiences even if you think they are very unlikely?". Belief flexibility is measured dichotomously with the creation of an alternate explanation rated as 'yes' or 'no'.

Another measure of belief flexibility is the Bias against Disconfirmatory Evidence (BADE; Moritz & Woodward, 2006). The BADE was specifically developed to examine a type of cognitive bias, whereby an individual's strongly held belief is maintained or revised when presented with disconfirmatory evidence. The BADE task consists of 16-30 delusion-neutral scenarios, for example, "Jenny can't fall asleep" along with four scenario-based sentences, for example, "Jenny is nervous about her exam the next day" (neutral lure); "Jenny is worried about her ill mother" (emotional lure); "Jenny loves her bed" (absurd); and "Jenny is excited about Christmas morning" (true). Of the four sentences, there is one true

interpretation (initially implausible but later plausible), two lure interpretations - emotional and neutral lures (initially plausible but later implausible), and one absurd interpretation (implausible throughout). The plausibility of the four interpretations is rated for each scenario on a scale of 0 (*poor*) to 10 (*excellent*).

Upon rating the statements, this procedure is repeated a further three times with additional information provided at each trial, for example, "Jenny can't fall asleep" (second trial); "Jenny can't wait until it is finally morning" (third trial); and "Jenny wonders how many presents she will find under the tree" (fourth trial). Upon reflecting on the information at each trial, individuals can adjust their ratings. When scoring the original version of the BADE task, reductions in the rating for the 'lure' item were evidence for the BADE (Moritz & Woodward, 2006). However, in the newer versions, results have been operationalised as evidence integration' (the degree to which information has been integrated) and 'conservatism' (reduced willingness to provide high plausibility ratings when justified; Speechley et al., 2012). It is of note, that over the years, there have been several versions of the BADE task, but they continue to share the same procedure.

Research on Belief Flexibility

The majority of studies on belief flexibility have been rooted in psychosis, with consistent associations with schizophrenia and the severity of delusional symptoms (Eisenacher & Zink, 2017; Garety et al., 2005). More recently, a study by Everaert et al. (2018) investigated interpretation bias and belief inflexibility with people experiencing common mental health difficulties. The results indicated that the severity of depression correlated with increased negative and decreased positive interpretation biases. However, the severity of social anxiety is significantly associated with an increased negative interpretation

bias. This would suggest that people experiencing greater severity of depression or social anxiety may be more biased and inflexible regarding their interpretations of situations.

Recent studies (Sanford et al., 2014; Speechley et al., 2012) have compared the performance on the BADE task between people experiencing high- and low-delusional schizophrenia with people experiencing obsessive-compulsive disorder or bipolar disorder, and healthy controls. Both studies found poorer integration of evidence was associated with the high delusional sample compared to both the low delusional group and the healthy controls.

Similarly, a study by Garety et al. (2013), investigated the differences in reasoning bias, as measured by the MADS and EoE, in relation to delusions subgroups (persecutory and grandiose delusions). The results indicated that the three variables measuring belief flexibility (PM, RTHC, and AE) were significantly different between the two groups. The grandiose delusions group showed an increased likelihood, compared to the persecutory group, of demonstrating reasoning bias.

Conversely, some studies have reported non-significant associations between delusions and evidence integration impairment (Eifler et al., 2014; Moritz et al., 2010; Veckenstedt et al., 2011) with even less evidence demonstrated for a positive association between delusions and evidence integration impairment (Riccaboni et al., 2012). However, it is of note that methodological limitations of the individual studies such as a lack of power in the sample size (Moritz et al., 2010) or a small variance of scores (Eifler et al., 2014) need to be considered in relation to the association not reaching significance.

A metanalytic review of 35 studies investigated the relationship between delusions in psychosis and cognitive biases, including the BADE. The sample consisted of patients with schizophrenia experiencing delusions, individuals with schizophrenia no longer experiencing

delusions, and healthy controls. The results indicated the group diagnosed with schizophrenia and experiencing delusions demonstrated increased BADE compared to the group with schizophrenia who historically experienced delusions. This would suggest BADE may be specifically related to delusions as opposed to schizophrenia alone (Mclean et al., 2016).

Research investigating belief flexibility in sub-clinical samples also found significant associations between delusions and belief flexibility (Menon et al., 2013; Zawadzki et al., 2012). A study by Orenes et al. (2012) examined whether BADE was evident in a non-clinical population with low and high schizotypal traits, as measured by the Schizotypal Personality Questionnaire (SPQ). The Orenes et al. (2012) study was an extension of the Moritz and Woodward (2006) study which found a significant association between BADE and people with schizophrenia. The results for Orenes et al. (2012) indicated a non-significant difference in BADE between the low and high SPQ. Accordingly, the findings are consistent with previous literature and may suggest that the BADE is demonstrated in clinical samples but not in non-clinical samples. However, it could also be argued that the detection of BADE may be less sensitive for non-clinical schizotypy as opposed to there being no difference between the low and high groups. Alternatively, it could be explained by the characteristics of schizotypy, for example, greater belief flexibility is observed in non-clinical schizotypy compared to people diagnosed with schizophrenia (Juarez-Ramos et al., 2014).

Rationale for the Review

There is a plethora of evidence demonstrating a clear association between belief flexibility and delusional severity in both clinical and non-clinical samples. However, little is known about the specific delusional symptoms, for example, paranoia and the relationship with belief flexibility. To the authors' knowledge, a systematic review examining the relationship between paranoia and belief flexibility has not been previously conducted.

Aim

The aim of the current review is to develop an understanding of the current literature on belief flexibility specifically regarding paranoia. The review will aim to address the following research questions, firstly, is there a relationship between belief flexibility and paranoia? Secondly, is there a difference in belief flexibility across the low and high paranoia groups?

Methods

The current review adhered to the statement of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009). This paper did not require ethical approval as it describes a literature review. The decision to complete a narrative synthesis was based on the limited number of studies available to perform a meta-analysis in the clinical and non-clinical subsections.

Registration of the Review Protocol

The review protocol was pre-registered and made public on the Open Science Framework (OSF) account at Sheffield University. The pre-registration link is https://osf.io/e5hsg/?view_only=cd8760c9ebce4cd9a0f7a7b54986383e

Search Strategy

A systematic literature search of three electronic databases: PsychINFO, MEDLINE, and Scopus was conducted. Three databases were used as this was considered to provide sufficient retrieval of results (Siddaway et al., 2019). Databases were searched using predefined search criteria using the Participant, Intervention, Control, Outcome (PICO) framework (Richardson et al., 1995) and where Medical Subject Headings (MeSH) were available, these terms were exploded and combined (Appendix A).

The search terms were developed by collating search terms used in previous related reviews and considering synonyms for the terms. To improve the search results assistance was gained from the librarian specialising in Psychology at Sheffield University.

The support included the use of Boolean operators (*AND* and *OR)* as well as the truncation technique to broaden the search terms to include different word endings and spellings, for example, using *Paranoi** rather than *Paranoia* and *Paranoid.*

The following search strings were used to search titles, abstracts, and keywords of publications with no limits placed on the database ("Paranoi*" OR "Persecut*" OR "Delusion*" OR "Suspici*") AND ("Belief Flexibility" OR "Belief Updating" OR "Belief Formation" OR "Bias Against Disconfirmatory Evidence" OR "BADE" OR "Evidence Integration" OR "Maudsley Assessment of Delusions Schedule" OR "MADS" OR "Explanation of Experiences Assessment" OR "EoE").

To aid the removal of duplications and enable hand searching of citations, articles were exported to the Mendeley software. Forward and backward manual citation searching was also conducted to identify additional articles from the reference lists within the studies included in the current review. This was to ensure that a thorough search of studies was completed for the current review.

Existing grey literature was not included in the review as a method to ensure a high standard of study quality (Aromataris & Ritano, 2014; Pappas & Williams, 2011). During instances when the studies appeared to assess the association between paranoia and belief flexibility but were not explicitly reported within the findings, the authors were contacted via email to gain or clarify the information.

Study Selection

Initial searching of articles yielded 501 papers. Articles were then extracted to Mendeley, and duplicate papers were removed from the search results ($n = 183$). Titles and abstracts of identified studies were screened by the author ($n = 318$) to determine relevance and subsequently checked against the inclusion and exclusion criteria. Articles that did not meet the inclusion criteria were excluded and a rationale was provided ($n = 228$). The full texts of the remaining papers ($n = 90$) were then screened. Additionally, hand searching of reference lists and studies citing the included article was carried out to gain a comprehensive search result. This process resulted in an additional study ($n = 1$).

To ensure the reliability of the selection process, an independent rater checked a proportion of the papers (15%; $n = 14$) during full-text screening. Inter-rater reliability showed substantial levels of agreement, Kappa = .63, $p < .01$, 95% CI (-.01,1.28). Discrepancies were resolved through discussion and/or contacting the original authors for further information.

The extracted data was organised on Microsoft Excel software using a review-specific form. Information extracted from the primary studies included primary author, year of publication, title, design, country, sample characteristics (n, age, gender), paranoia and belief flexibility measure, key findings, and the data quality score (Table 1).

Eligibility criteria

For the review, the eligible sample was found in the studies that reported findings on an adult (18+) clinical or subclinical sample. The rationale to be inclusive of these participants was based on the current aims of the review and to develop the current understanding of belief flexibility and levels of paranoia as the findings are inconclusive.

Inclusion Criteria

Studies were included based on the following criteria: (1) participants were adults, male, female, and non-binary, aged 18+ years old, (2) the sample included clinical and/or subclinical paranoid individuals, (3) quantitative studies assessing paranoia and belief flexibility using validated measures, (4) quantitative studies reporting statistical results on paranoia and belief flexibility, (5) RCTs reporting baseline data, (6) mixed-methods studies reporting quantitative data relevant to the review aims, (7) articles using quantitative non-randomised controlled trials (including cohort studies, case-control studies, cross-sectional studies), and (8) papers using quantitative case studies and series.

Exclusion Criteria

Studies were omitted based on the following criteria: (1) full text was not available in English, (2) retrieved articles were qualitative studies including narrative reviews, literature reviews, editorials, dissertations, book chapters and commentaries, (3) studies that focused on children and young people below the age of 18 years old, and (4) grey literature.

Quality Appraisal

To evaluate the quality of the included studies, critically appraise the study findings, and eliminate the risk of bias, the Effective Public Health Practice Project (EPHPP) Quality Assessment Tool for Quantitative Studies (Thomas et al., 2004: Appendix B) was used.

The EPHPP is commonly used within healthcare-related systematic reviews as there is clear guidance regarding use and the tool has good content and construct validity (Thomas et al. 2004; Appendix B) as well as good inter-rater reliability (Armijo-Olivio et al., 2012). Moreover, based on a review by Deeks et al. (2003) the EPHPP is considered one of the "best tools" for assessing study quality.

The EPHPP evaluates eight components (1) selection bias, (2) study design, (3) confounders, (4) blinding, (5) data collection methods, (6) withdrawals and dropouts, (7) intervention integrity, and (8) analysis, with a quality rating for each component denoted by a numerical value (1 = strong, 2 = moderate, and 3 = weak). Overall global rating is calculated by summing the individual scores for the eight components. Classification for overall global ratings is strong (no weak ratings), moderate (one weak rating) or weak (two or more weak ratings).

Within the current review, ratings of the study were lower, for example, if the sample did not represent a clinical or subclinical paranoid population, the study design was not indicated, confounding variables including age and sex were not controlled, blinding was not carried out, validated measures of paranoia and the use a belief flexibility tasks, for example, Green's Paranoia Thought Scale-B (Green et al., 2008) and Bias Against Disconfirmatory Evidence (Moritz & Woodward, 2006), respectively, were not used, withdrawal and dropout rates were not stated, and key decisional information was omitted for the intervention and analysis components, for example, the rationale for the p value used in the study.

In addition to the global rating scores of the paper, the component rating scores were also reported. This was to facilitate transparency and mitigate the likelihood of misleading readers when only reporting global ratings (Liberati et al., 2009).

A second independent rater also critically appraised a percentage of the included papers to aid reliability (37.5% ; $n = 3$). Disagreements in component and global ratings between the raters were resolved through further discussion until consensus was reached. The discrepancy in ratings was identified for the study sample representing the target population and whether confounding variables were sufficiently attended to in the analysis. Inter-rater reliability was calculated using Kappa and interpreted following the criteria proposed by

Landis and Koch (1977). There was substantial agreement between the raters, Kappa = .73, *p* < .001, 95% CI (.52, .94).

Results

Summary of Studies

A total of eight studies were included in the systematic review. The PRISMA flow diagram outlines the different phases of the review (Figure 1; Moher et al., 2009). All articles were peer-reviewed and published between January 1972 and February 2022. It is of note that the earliest date reflects the earliest found study rather than a predefined date. A summary of each study's characteristics and key findings can be found in Table 1.

The studies included in the review predominantly employed a cross-sectional research design (*n* = 6) to assess the relationship between paranoia and belief flexibility. Two studies did not report the research design (Lavigne et al., 2020; Woodward et al. 2007). The majority of studies in the review were carried out in Canada (*n* = 4) followed by the United States of America (*n* = 2), China (*n* = 1), and Germany (*n* =1).

Across the studies, study samples were recruited via local Universities (*n* = 3), Amazon's Mechanical Turk - Online platform (*n* = 2), Central Institute of Mental Health (*n* = 1), and psychiatric hospitals (*n* = 1). One study did not report recruitment information (Lavigne et al., 2020).

The total number of participants included across the studies was 906. However, two studies (Buchy et al., 2007; Woodward et al., 2007) used the same participants for the study samples and therefore have overlapping samples. For the purpose of the review, the same participants will be considered twice for the narrative synthesis.

The review sample included male ($n = 374$), female ($n = 419$), and non-binary ($n = 1$) participants. Two studies did not report the gender of the samples thus, accounting for 112 participants in the review (Woodward et al., 2006, 2007). The average age of the sample across six studies was 33.99 years. One study did not report on age (Woodward et al., 2006), whilst one study reported age within age brackets and the *n* (Bronstein et al., 2019; see Table 1 for *n* values).

Of the eight studies, six studies consisted of non-clinical samples (Bronstein et al., 2019; Buchy et al., 2007; Deng et al., 2022; Lavigne et al., 2020; Woodward et al., 2007; Zawadzki et al., 2012; $n = 805$) and two studies consisted of clinical samples with non-clinical samples recruited as controls (Eifler et al., 2014; Woodward et al., 2006; $n = 101$, please see Table 1 for sample breakdown).

Of the clinical samples of interest, diagnoses included schizophrenia ($n = 29$) and schizophrenia spectrum disorders ($n = 14$) (Eifler et al., 2014; Woodward et al., 2006, respectively, see Table 1).

Despite efforts to identify studies investigating differences in belief flexibility between low and high paranoia groups, only three studies were found to include a sample consisting of these groups (Buchy et al., 2007; Woodward et al., 2006; Zawadzki et al., 2012). In Buchy et al. (2007) and Woodward et al. (2006) study, participants scoring greater than the 90th percentile on the Schizotypal Personality Questionnaire (SPQ; Raine, 1991) formed the high group ($n = 74$) whilst participants scoring less than the 10th percentile on the SPQ formed the low group ($n = 64$). In Zawadzki et al. (2012) study participants were differentiated using a median split of the Peters et al. Delusions Inventory – Paranoia item score (PDI; Peters et al. 1999) for the low ($n = 58$) and high ($n = 59$) groups.

Across the studies, paranoia was measured using the Positive and Negative Syndrome Scale -Item 6 (PANSS; Kay & Opler, 1987; $n = 2$), SPQ (Raine, 1991; $n = 3$), Green's Paranoia Thought Scale (GPTS; Green et al., 2008; $n = 1$), Revised GPTS (R-GPTS; Freeman et al., 2019: $n = 1$), and PDI (Peters et al. 1999; $n = 1$). To measure belief flexibility, all the studies used a version of the bias against disconfirmatory evidence task (BADE; Moritz & Woodward, 2006).

Figure 1

PRISMA Diagram
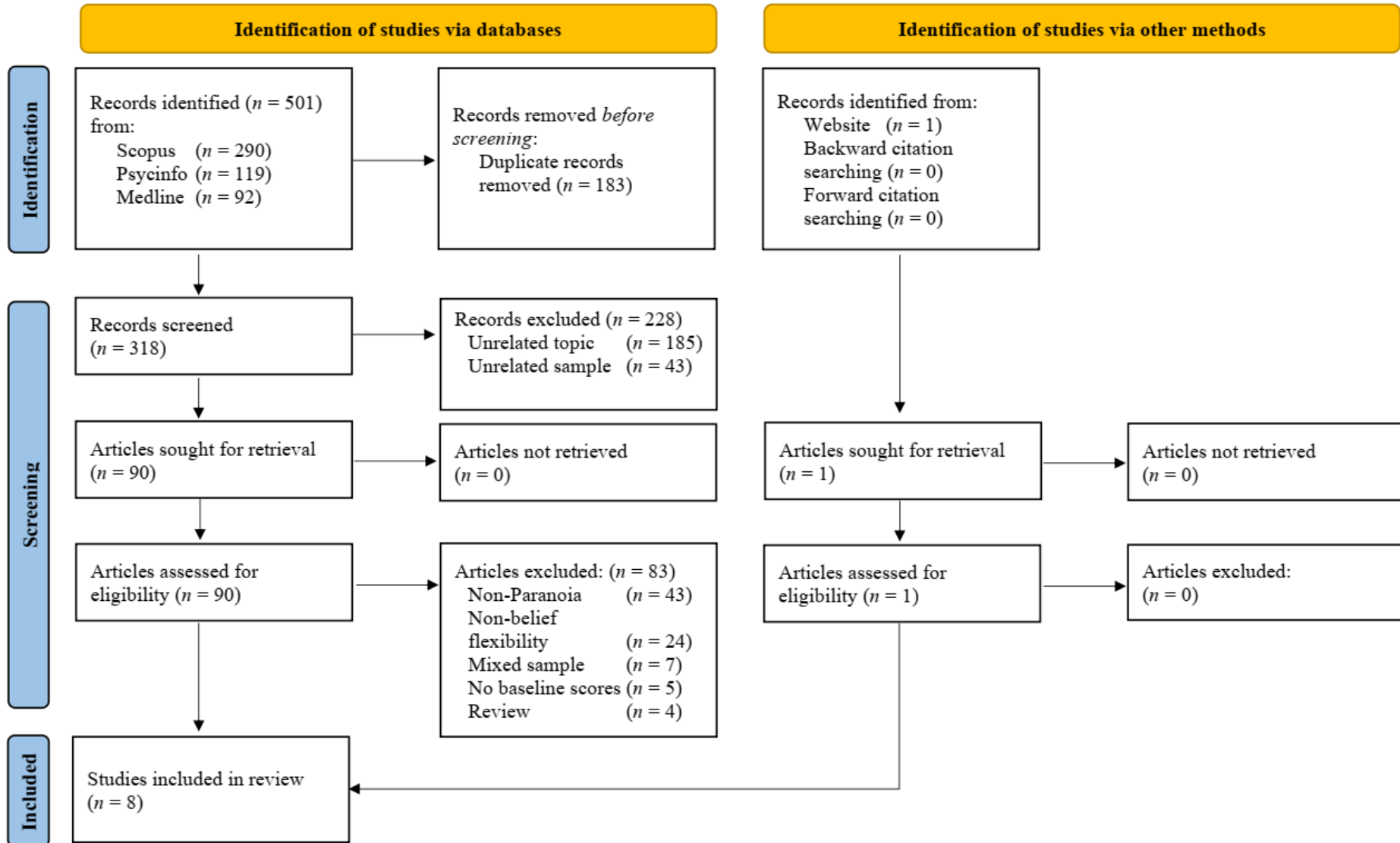
Table 1

Study Characteristics for the Eight Studies Included in the Review

					Measures			
Source (year)	Title	Design	Country Setting	Sample Characteristics *n* age (years) gender (*n*)	Paranoia	Belief Flexibility	Key Findings	Data Quality
Non-Clinical Sample								
Bronstein et al. (2019)	Pathways to paranoia: Analytic thinking and belief flexibility	CS	United States Amazon's Mechanical Turk - Online Crowdsource Platform	Non-Clinical 231 18 - 29 (*n* = 62) 30 - 39 (*n* = 84) 40 - 49 (*n* = 39) 50+ (*n* = 46) F = 130; M = 101	GPTS-B	BADE	**EII and paranoia** β = .46, $t(230)$ = 8.00, $p < .05$ CI = .35-.57 **PII and paranoia** β = .47, $t(230)$ = 8.04, $p < .05$ CI = .35-.57 **NII and paranoia** β = .47, $t(230)$ = 8.11, $p < .05$ CI = .35-.57	3
Deng et al. (2022)	Developing a novel assessment of interpretation flexibility: Reliability, validity and clinical implications	CS	United States Amazon's Mechanical Turk - Online Crowdsource Platform	Non-Clinical 274 *M* = 37.77; *SD* = 10.83 F = 130 ; M = 143 Non-Binary = 1	R-GPTS	Emotional BADE	Negative Outcome Scenarios **PIB and paranoia** β = -.13, $p = .63$ **NIB and paranoia** β = .28, $p < .001$ Positive Outcome Scenarios **PIB and paranoia** β = 0.04, $p = .4)$ **NIB and paranoia** β = 0.20, $p < .001$	3

					Measures			
Source (year)	Title	Design	Country Setting	Sample Characteristics *n* age (years) gender (*n*)	Paranoia	Belief Flexibility	Key Findings	Data Quality
Lavigne et al. (2020)	Functional brain networks underlying evidence integration and delusional ideation	NS	Canada Setting NS	Non-Clinical 41 *M* = 35.44; *SD* = 12.59) F = 22 ; M = 19	SPQ suspicious subscale	BADE - Behavioural	**SPQ suspiciousness and behavioural BADE** *β* = -0.61, *p* < .05	3

Clinical with Non-Clinical Control Sample

Eifler et al. (2014)	Neurocognitive capabilities modulate the integration of evidence in schizophrenia	CS	Germany Schizophrenia – Central Institute of Mental Health in Mannheim, Germany. Healthy Controls – Unclear	Schizophrenia *n* = 29 *M* = 37.14; *SD* = 10.05 F = 7 ; M = 22 Healthy Controls *n* = 29 *M* = 34.76; *SD* = 11.12 F = 9 ; M = 20	PANSS P6	BADE	**Disconfirmatory evidence** *r* = 0.15, *p* = .27 **Confirmatory evidence** *r* = 0.02, *p* = .89	3

					Measures			
Source (year)	Title	Design	Country Setting	Sample Characteristics *n* age (years) gender (*n*)	Paranoia	Belief Flexibility	Key Findings	Data Quality
Woodward et al. (2006)	The contribution of a cognitive bias against disconfirmatory evidence to delusions: A study in an Asian sample with first episode schizophrenia spectrum disorders	CS	Hong Kong, China. Two psychiatric Hospitals	First episode Schizophrenia and Non-Clinical *n* = 43 Healthy Control *n* = 17 Delusional *n* = 14 Non-Delusional *n* = NS *M* = NS; *SD* = NS F = NS ; M = NS	PANSS P6	BADE	**Delusional vs. non-delusional patients;** *r* = .31, *p* < .05 **Delusional patients vs. controls**. *r* = .36, *p* = .11	2
Non-Clinical Sample with High and Low Groups								
Buchy et al. (2007)	A cognitive bias against disconfirmatory evidence (BADE) is associated with schizotypy	CS	Canada University	Non-Clinical Students High Schizotypy 37 *M* = 19.11; *SD* = 1.83 F= 23; M = 14 Low Schizotypy 32 *M* = 19.67; *SD* = 2.28 F = 20 ; M = 12	SPQ suspicious subscale	BADE	**SPQ-suspiciousness and BADE** *r* = -.22, *p* >.06	2

					Measures			
Source (year)	Title	Design	Country	Sample Characteristics	Paranoia	Belief Flexibility	Key Findings	Data Quality
			Setting	n age (years) gender (n)				
Woodward et al. (2007)	A bias against disconfirmatory evidence Is associated with delusion proneness in a nonclinical sample	NS	Canada University	Non-Clinical undergraduate students 69 High non-clinical schizotypal $n = 37$ $M = 19.11$; $SD = 1.83$ F= 23 ; M = 14 Low non-clinical schizotypal $n = 32$ $M = 19.67$; $SD = 2.28$ F = 20 ; M = 12	SPQ suspicious subscale	BADE	**SPQ suspicious subscale and BADE initial belief** (rating after sentence 1 and after sentence 2) $r = -.02, p > .05$ **SPQ suspicious subscale and BADE evidence integration** (rating after sentence 2 and after sentence 3) $r = .12, p > .05$	3

				Measures				
Source (year)	Title	Design	Country Setting	Sample Characteristics *n* age (years) gender (*n*)	Paranoia	Belief Flexibility	Key Findings	Data Quality
Zawadzki et al. (2012)	Cognitive factors associated with subclinical delusional ideation in the general population	CS	Toronto, Canada University	Non-clinical undergraduate students 121 *M* = 30.7 ; *SD* = 12.6 F = 78 ; M = 43 High PDI *n* = 58 Low PDI *n* = 59	PDI-Paranoia	BADE	**PDI paranoia and BADE ratio** (reduction in confidence between 1st and 3rd trial for lure interpretations as a ratio as an increased confidence for the true interpretations) *r* = -.04, *p* > .05 **PDI and Liberal Selection within BADE** (number of lure or absurd interpretations endorsed) *r* = .05, *p* > .05 **High PDI Evidence integration** *r* = .18, *p* >.05 **Low PDI Evidence integration** *r* = .07, *p* >.05	3

Note. BADE = Bias against disconfirmatory evidence, CS = Cross-sectional, EII = Evidence Integration Impairment F = Female. GPTS-B =

Green's paranoia thought scale – B. M = Male. NIB = Negative integration bias. NII = Negative Integration Impairment. NS = Not specified.

PANSS = Positive and negative syndrome scale. PDI = Peters et al. delusions inventory. PIB = Positive integration bias. PII = Positive

Integration Impairment. R-GPTS = Revised - Green's paranoia thought scale. SPQ = Schizotypal personality questionnaire.

Study Quality

Quality appraisals of the studies included in the review are presented in Table 2. The risk of selection bias across the eight studies was rated as 'strong' for three studies (Bronstein et al., 2019; Deng et al., 2022; Eifler et al., 2014), as 'moderate' for three studies (Buchy et al., 2007; Lavigne et al., 2020; Woodward et al., 2006), and as 'weak' for two studies (Woodward et al., 2007; Zawadzki et al., 2012). Selection bias was evident in the 'moderate' and 'weak' rated studies due to limited information about the recruitment procedure, convenience sampling and/or the sample was not representative of the target population.

Regarding study design, all studies were rated 'moderate'. Most studies did not control for confounding variables ($n = 4$) and were therefore rated 'weak' (Bronstein et al., 2019; Deng et al., 2022; Eifler et al., 2014; Lavigne et al., 2020). The 'moderate' rated studies ($n = 3$) controlled for 60%-79% of the confounding variables (Buchy et al., 2007; Woodward et al., 2007; Zawadzki et al., 2012) whilst the 'strong' rated study (Woodward et al., 2006) controlled for 80-100% of the confounding variables. Common confounding variables included age, sex, education, and race. All studies were rated 'weak' for blinding as both the participants and the researchers were aware of the purpose of the study.

Data collection and data analysis across the studies were rated 'strong'. Across the studies, measurements of paranoia and belief flexibility were completed using validated and reliable tools. Data analysis was also appropriate for the research questions and the research design. Of the eight studies, four studies reported withdrawal and/or dropout rates and therefore, rated 'strong'(Buchy et al., 2007; Deng et al., 2022; Eifler et al., 2014; Zawadzki et al., 2012). Three studies were rated 'moderate' due to limited transparency of withdrawal and/or dropout rates (Bronstein et al., 2019; Lavigne et al., 2020; Woodward et al., 2006). There was no reference to the withdrawal and/or dropout in the Woodward et al. (2007) study

and therefore the study was rated 'weak'. Intervention integrity was not rated as it was not applicable to the current review. Overall, the global rating for most studies was 'weak' ($n = 6$; Bronstein et al., 2019; Deng et al., 2022; Eifler et al., 2014; Lavigne et al., 2020; Woodward et al., 2007; Zawadzki et al., 2012) whilst two studies were rated 'moderate' (Buchy et al., 2007; Woodward et al., 2006).

Table 2

Quality Assessment Scores for the Included Studies

Study	Selection Bias	Study Design	Confounder	Blinding	Data Collection	Withdrawal and Dropout	Intervention Integrity	Data Analysis	Global Rating
Bronstein et al. 2019)	Strong	Moderate	Weak	Weak	Strong	Moderate	N/A	Strong	Weak
Buchy et al. (2007)	Moderate	Moderate	Moderate	Weak	Strong	Strong	N/A	Strong	Moderate
Deng et al. (2022)	Strong	Moderate	Weak	Weak	Strong	Strong	N/A	Strong	Weak
Eifler et al. (2014)	Strong	Moderate	Weak	Weak	Strong	Strong	N/A	Strong	Weak
Lavigne et al. (2020)	Moderate	Moderate	Weak	Weak	Strong	Moderate	N/A	Strong	Weak
Woodward et al. (2006)	Moderate	Moderate	Strong	Weak	Strong	Moderate	N/A	Strong	Moderate
Woodward et al. (2007)	Weak	Moderate	Moderate	Weak	Strong	Weak	N/A	Strong	Weak
Zawadzki et al. 2012)	Weak	Moderate	Moderate	Weak	Strong	Strong	N/A	Strong	Weak

Note. N/A = Not applicable.

Relationship between Paranoia and Belief Flexibility

All studies compared the relationship between paranoia and belief flexibility. To aid understanding and the interpretation of the results from the included studies, the narrative synthesis will focus on the non-clinical sample and then the clinical sample with the non-clinical control group.

Non-Clinical Sample

Three studies reported the association between paranoia and belief flexibility in the non-clinical sample. Of the three studies, two studies examined the associations between the variables SPQ-Suspiciousness item and behavioural BADE (Lavigne et al., 2020) and GPTS and BADE positive and negative interpretation bias (Deng et al., 2022). However, in Bronstein et al. (2019) the relationship between paranoia, as assessed by the R-GPTS, and BADE was examined, and so the narrative synthesis will follow this format.

A negative correlation was found between SPQ-suspiciousness and behavioural BADE (β = -.61, p < .05; Lavigne et al., 2020). The Deng et al. (2022) study used two conditions (negative and positive) with two emotional BADE components (positive interpretation bias and negative interpretation bias) to assess belief flexibility. The negative outcome condition consisted of sentences that initially encouraged positive interpretation bias but with a scenario outcome that was negative, therefore the disconfirming evidence in the task was the positive interpretation bias. The positive outcome condition consisted of sentences that initially encouraged negative interpretation bias but with a scenario outcome that was positive, therefore the disconfirming evidence in the task was the negative interpretation bias.

Within the negative outcome condition, a highly significant negative association was found between paranoia and NIB (β = .28, p < .001) but not PIB (β = -.13, p = .63). Within

the positive outcome condition, a highly significant negative association was found between paranoia and NIB ($\beta = .20$, $p < .001$) but not PIB($\beta = .04$, $p = .477$). The findings from both conditions would suggest that paranoia is consistently related to interpretation bias in the negative condition but inconsistently related to paranoia in the positive condition as the preference was for NIB. This would suggest that belief flexibility is reduced in paranoia, and this is more evident when presented with negative information.

Overall, within the non-clinical samples, the results appear mixed regarding the association between paranoia and various BADE tasks. The main difference between the two studies was the measurement of belief flexibility within BADE. One study looked at behavioural BADE whilst the other study analysed BADE positive and negative interpretation bias within the positive and negative conditions (Lavigne et al., 2020; Deng et al., 2022, respectively). Consequently, conclusions for the non-clinical sample are difficult to ascertain.

Despite the review finding mixed results for the association between paranoia and various BADE tasks, Bronstein et al. (2019) found evidence integration impairment, as measured by BADE, predicted paranoia ($\beta = .46$, $t(230) = 8.00$, $p < .05$). Moreover, positive, and negative integration impairment also predicted paranoia ($\beta = .47$, $t(230) = 8.04$, $p < .05$ and $\beta = .47$, $t(230) = 8.11$, $p < .05$, respectively). This may suggest that specifically BADE negative interpretation bias has a direct effect on paranoia. However, as there was only one study included in the review assessing the predictive relationship between the variables (Bronstein et al., 2019), the findings need to be interpreted with caution for the non-clinical sample.

Clinical Sample with a Non-Clinical Control Group.

Two studies used a clinical sample with a non-clinical sample as a control group (Eifler et al., 2014; Woodward et al., 2006). The two studies examining the PANSS P6 items with BADE reported different associations and will therefore be summarised separately. In Eifler et al. (2014) the correlational analysis was part of the secondary analysis and so only the results for the clinical sample were reported. The results indicated a non-significant association for PANSS P6 and BADE Disconfirmatory evidence and BADE Confirmatory evidence ($r = 0.15$, $p = .27$ and $r = 0.02$, $p = .89$, respectively). In Woodward et al. (2006) study BADE disconfirmatory evidence was examined with delusion, non-delusional, and control groups. A positive association was found for the delusion group when compared with the non-delusional group ($r = .31$, $p < .05$), but there was a non-significant association for the delusional group when compared with the control group ($r = .36$, $p = .11$). Consequently, the results from both studies would indicate paranoia when measured by PANNS P6 does not significantly correlate with belief flexibility when comparing clinical samples to non-clinical samples.

Relationship between Paranoia and Belief Flexibility in **Low** *and High Groups*

Three studies differentiated the sample into low and high non-clinical groups (Buchy et al., 2007; Woodward et al., 2007; Zawadzki et al., 2012). Despite differentiating groups into low and high groups, Buchy et al. (2007) and Woodward et al. (2007) only reported the combined (total sample) correlations for suspiciousness and belief flexibility. Buchy et al. (2007) found a non-significant correlation between suspiciousness and belief flexibility. Similarly, Woodward et al. (2007) study found a non-significant association between SPQ-suspiciousness and BADE initial belief (rating after sentence 1 and after sentence 2) and BADE evidence integration (rating after sentence 2 and after sentence 3). All *p's* $>.05$.

On the other hand, in Zawadzki et al. (2012) study, the results were reported for the low and high PDI groups. Nevertheless, the association between the evidence integration and paranoia in the low and high groups was non-significant (all *p's* >. 05). This would suggest both SPQ and PDI do not significantly correlate with belief flexibility. Although, Zawadzki et al. (2012) found a negative association between PDI and BADE ratio (an increase in the 'true' interpretation without the reduction in confidence for 'lure' interpretations) as well as PDI and liberal selection (number of endorsed lure or absurd interpretations) across the sample ($r = -.22$, $p < .05$ and $r = -.22$, $p < .0001$, respectively). This would suggest participants showed limited confidence in reducing the scores for the 'lure' item even when presented with disconfirmatory evidence. Moreover, there was a selection bias for the decision to choose a lure or an absurd item rather than, the highest rating for lure or absurd items for the BADE scenarios.

Discussion

Summary of Findings

The aims of the current systematic review were to assess the relationship between belief flexibility and paranoia and whether there was a difference in belief flexibility between people with low and high paranoia. A total of eight studies, yielding a sample size of 906, met the inclusion criteria for the review. All studies compared the relationship between paranoia and belief flexibility.

Non-Clinical Sample

Two studies reported on the association between paranoia and belief flexibility in the non-clinical sample, whilst one study examined the relationship between the variables. However, measurement of paranoia severity differed across the studies. One study used the SPQ-Suspiciousness item (Lavigne et al., 2020) whilst the other two studies used a variant of

the GPTS measure (Bronstein et al., 2019; Deng et al., 2022) to measure paranoia.

Additionally, across the three studies, belief flexibility was measured and reported using

various versions of the BADE, for example, behavioural BADE and emotional BADE. As a

result, there was limited opportunity to compare the findings and draw an overall conclusion

for this sample. However, Deng et al. (2022), found an association between BADE evidence

integration bias and paranoia. This was further supported by the Bronstein et al. (2019) study

which found negative evidence integration impairment predicted paranoia. As a result, the

findings suggest that a reduced ability in belief flexibility, specifically when information is

negative, may increase the likelihood of experiencing paranoia.

Clinical Sample with a Non-Clinical Control Group

Two studies used a clinical sample with a non-clinical sample as a control group

(Eifler et al., 2014; Woodward et al., 2006) Both studies found a non-significant association

for disconfirmatory evidence in the clinical sample compared to the non-clinical sample even

with different sample sizes. It can therefore be argued that the non-significant association

occurred irrespective of Woodward et al. (2006) using a larger sample size, which can be

claimed to increase the likelihood of a significant result. However, as both studies did not

report a power analysis, findings could be subject to a type one or type two error (Field

2009).

Comparative analysis, using a non-clinical sample was not completed in the Eifler et

al. (2014) study as this was not of interest to the authors. However, in the Woodward et al.

(2006) study the delusional sample was compared to the non-delusional sample regarding

paranoia and BADE disconfirmatory evidence, which was found to be significant.

Consequently, suggesting a relationship between BADE disconfirmatory evidence is

associated with the delusional sample when compared to the non-delusional sample but not

when compared to the non-clinical sample. These findings are inconsistent with expectations, as BADE is expected to be associated with paranoia (Woodward et al., 2007).

Relationship between Paranoia and Belief Flexibility in Low and High Groups

Three studies differentiated the sample into low and high sub-clinical groups (Buchy et al., 2007; Woodward et al., 2007; Zawadzki et al., 2012). Both, Buchy et al., 2007 and Woodward et al. (2007), found a non-significant association between paranoia and belief flexibility. However, the results from these studies should be interpreted with caution as the correlational analysis was performed for the whole sample and not the subgroups. As a result, it could be argued that by pooling the data from both low and high groups, the data would not be normally distributed, as the data set only includes extreme values from either side of the distribution curve, therefore, reducing the reliability of the results.

Conversely, the results were reported for the low and high PDI groups in Zawadzki et al. (2012) study. Nevertheless, the association between the evidence integration and paranoia in both groups was non-significant. This would suggest that belief flexibility when measured by BADE does not correlate to the SPQ and PDI measures of paranoia within the low and high groups. However, as the findings between the low and high groups are based on one study, further studies are warranted to draw conclusions with confidence.

Strengths and Limitations

Previous reviews have examined belief flexibility in relation to delusions in psychosis, schizophrenia, and over the course of psychosis, see reviews by Mclean et al., (2017), Moritz et al. (2014), and Eisenacher and Zink. (2017), respectively. However, to the authors' knowledge, the current review was the first to assess the relationship between belief flexibility and paranoia as well as assess the current understanding of the differences in belief flexibility between the low and high paranoia groups.

The current review can be commended for utilising a systematic approach following the Prisma checklist to identify and review the included studies forming the review. This approach was used to aid methodological rigour to increase the reliability and validity of the review. Moreover, inter-rater reliability checks were completed during the screening phase and during the quality assessment of the included studies, which yielded substantial inter-rater reliability at both times (Landis & Koch, 1977). By including multiple raters at various points of the review it increased the methodological integrity of the review and ensured consistency and consensus of the ratings and the included papers.

However, the understanding gained from the review, specifically the association between belief flexibility and paranoia, is limited, due to several shortcomings of the included studies. Firstly, the number of published studies included in the review was limited. Only peer-reviewed papers were included due to the assumption that these papers are of high quality. However, by excluding literature, as the studies were not peer-reviewed (grey literature) or papers that were not published in English due to translation difficulties, the current review may have overlooked relevant studies. For example, studies with null findings may have provided a further understanding of belief flexibility and paranoia (Rosenthal, 1979).

Secondly, across the studies, belief flexibility was measured using various iterations of the BADE task, for example, behavioural, emotional and the original BADE task. As a result, for the most part, a comprehensive synthesis of the findings was difficult due to the use of different measures and inconsistent results, therefore, it was difficult to draw strong conclusions from the review. Similarly, paranoia was typically measured using a single item from a standardised measure, for example, SPQ-Suspiciousness or the PANSS P6 item, with two studies measuring paranoia severity using the GPTS or the R-GPTS. It is therefore difficult to say whether the measure itself impacted the results yielded in studies and

subsequently in the review. For example, when using the SPQ-Suspiciousness item a non-significant association for BADE was observed. However, the study using the GPTS measure, which assesses paranoia across several items, found a significant association with BADE. It may be that a single-item measure does not encompass the complexity of the formation and maintenance of paranoia, therefore, the validity of these studies can be questioned (Garety et al., 2001; Freeman & Garety, 2004).

Finally, the overall rating of quality for most of the studies was 'weak' with two studies rated as 'moderate'. Ratings of 'weak' and 'moderate' were due to sampling bias or the study not attending to confounding variables. As a result, the conclusions drawn from the current review should be interpreted with caution.

Research Limitations, Clinical Implications, and Future Research

Evidently, based on the eight studies included in the review, there is limited research and understanding of belief flexibility and paranoia. Previous research has included people experiencing paranoia within a sample of people experiencing delusions, however, the studies did not differentiate the findings for the subgroups and therefore they could not be included in the review, for example, So et al. (2012). Moreover, the conclusions drawn from the study cannot be applied to people experiencing a specific sub-type of delusion, for example, persecutory delusions or grandiose delusions. Future studies should assess the subtypes of delusions individually to allow for specific clinical application.

As evidenced in the current review, paranoia is prevalent in clinical and non-clinical samples. Whilst the review findings cannot with confidence say there is a strong association between belief flexibility and paranoia due to limitations within the studies. The findings do suggest there may be a possible link. As a result, therapeutic work could focus on the cognitive element of Cognitive Behaviour Therapy (CBT) to support belief flexibility in

relation to paranoid ideation. Freeman et al. (2016) noted six key self-help steps in reducing paranoia using CBT, with the third element focussing on reviewing the paranoid thought and attempting to find alternate explanations, thus updating the individual's belief about the paranoid thought.

The studies included in the review were predominantly cross-sectional in design, therefore, leading to difficulties in ascertaining causal inferences between belief flexibility and paranoia. To overcome the design limitation, future studies could follow a longitudinal design and assess both clinical and non-clinical populations experiencing paranoia. This may support the understanding of belief flexibility across time and how best to intervene regarding the management of paranoia.

Additionally, as per the limitations section in the current review, the global rating for most of the included studies was 'weak' ($n = 6$) with two studies rated moderate, future studies should address the methodological shortcomings by attending to, for example, confounding variables, as a method of creating strong quality research. A replication of the current review and conducting a metanalytic review, when high-quality studies are available, is also warranted in the future. This is to aid understanding of paranoia and belief flexibility and to increase confidence in the conclusions drawn from the review.

Conclusions

In conclusion, the current review found mixed findings for the relationship between belief flexibility and paranoia which were evident across both clinical and non-clinical samples. Interestingly, when paranoia was measured using a paranoia-specific measure compared to a single item from a standardised measure, the association between paranoia and the BADE evidence integration bias was significant. Consistent with this finding, one study also found BADE evidence integration impairment predicted paranoia. This would suggest

that there appears to be a relationship between paranoia and belief flexibility, however, due to the limited studies in the review and their findings, the result from the current review should be interpreted with caution. Future studies could attend to the methodological shortcomings of the current studies included in the review to aid the understanding of belief flexibility in people experiencing paranoia.

References

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2012).
Assessment of study quality for systematic reviews: a comparison of the Cochrane
Collaboration Risk of Bias Tool and the Effective Public Health Practice Project
Quality Assessment Tool: Methodological Research. *Journal of Evaluation in
Clinical Practice*, *18*(1), 12-18. https://doi.org/10.1111/j.1365-2753.2010.01516.x

Aromataris, E., & Riitano, D. (2014). Constructing a search strategy and searching for
evidence. *American Journal of Nursing*, *114*(5), 49-56.
doi: 10.1097/01.NAJ.0000446779.99522.f6

Bronstein, M. V., Everaert, J., Castro, A., Joormann, J., & Cannon, T. D. (2019). Pathways to
paranoia: Analytic thinking and belief flexibility. *Behaviour Research and
Therapy*, *113*, 18-24. https://doi.org/10.1016/j.brat.2018.12.006

Buchy, L., Woodward, T. S., & Liotti, M. (2007). A cognitive bias against disconfirmatory
evidence (BADE) is associated with schizotypy. *Schizophrenia Research*, *90*(1-3),
334-337. https://doi.org/10.1016/j.schres.2006.11.012

Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovitch, C., Song, F., ... & Altman,
D. G. (2003). Evaluating non-randomised intervention studies. *Health technology
assessment (Winchester, England)*, *7*(27), iii-173. https://doi.org/10.3310/hta7270

Deng, W., Everaert, J., Creighton, M., Bronstein, M. V., Cannon, T., & Joormann, J. (2022).
Developing a novel assessment of interpretation flexibility: Reliability, validity and
clinical implications. *Personality and Individual Differences*, *190*, 111548.
https://doi.org/10.1016/j.paid.2022.111548

Eifler, S., Rausch, F., Schirmbeck, F., Veckenstedt, R., Englisch, S., Meyer-Lindenberg, A.,
... & Zink, M. (2014). Neurocognitive capabilities modulate the integration of
evidence in schizophrenia. *Psychiatry Research*, *219*(1), 72-78.
https://doi.org/10.1016/j.psychres.2014.04.056

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

Everaert, J., Bronstein, M. V., Cannon, T. D., & Joormann, J. (2018). Looking through tinted
glasses: Depression and social anxiety are related to both interpretation biases and
inflexible negative interpretations. *Clinical Psychological Science*, *6*(4), 517-528.
https://doi.org/10.1177/2167702617747968

Field, A. (2009). *Discovering Statistics Using SPSS*, *Third Edition*. SAGE.

Fischoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective.
Psychological Review, *90*(3), 239–260. doi:10.1037/0033-295X.90.3.239

Freeman, D., Bradley, J., Waite, F., Sheaves, B., DeWeever, N., Bourke, E., ... & Dunn, G.
(2016). Targeting recovery in persistent persecutory delusions: a proof of principle
study of a new translational psychological treatment (the Feeling Safe
Programme). *Behavioural and Cognitive Psychotherapy*, *44*(5), 539-552.
doi:10.1017/S1352465816000060

Freeman, D., & Garety, P. A. (2004). *Paranoia: The psychology of persecutory delusions*.
Psychology Press.

Freeman, D., Garety, P. A., Fowler, D., Kuipers, E., Bebbington, P. E., & Dunn, G. (2004).
Why do people with delusions fail to choose more realistic explanations for their
experiences? An empirical investigation. *Journal of consulting and clinical
psychology*, *72*(4), 671. doi: 10.1037/0022-006X.72.4.671

Freeman, D., Loe, B., Kingdon, D., Startup, H., Molodynski, A., Rosebrock, L., . . . Bird, J.

    (2021). The revised Green et al., Paranoid Thoughts Scale (R-GPTS): Psychometric

    properties, severity ranges, and clinical cut-offs. *Psychological Medicine, 51*(2), 244-

    253. doi:10.1017/S0033291719003155

Garety, P. A., Freeman, D., Jolley, S., Dunn, G., Bebbington, P. E., Fowler, D. G., &

    Dudley, R. (2005). Reasoning, emotions, and delusional conviction in psychosis.

    Journal of Abnormal Psychology, *114*, 373–384. doi:10.1037/0021-843X.114.3.373

Garety, P. A., Gittins, M., Jolley, S., Bebbington, P., Dunn, G., Kuipers, E., ... & 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

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A

    cognitive model of the positive symptoms of psychosis. Psychological Medicine, *31*,

    189 –195. doi:10.1017/S0033291701003312

Garety, P. A., Ward, T., & Rus-Calafell, M. (2020). Beyond Belief—New Approaches to the

    Treatment of Paranoia. In J. Badcock & G. Paulik (Eds.), *A Clinical*

    Introduction to Psychosis (pp. 591-613). Academic Press.

Green, C. E. L., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., & Garety, P.

    A. (2008). Measuring ideas of persecution and social reference: The Green et al.

    Paranoid Thought Scales (GPTS). *Psychological Medicine*, *38*(1), 101-111.

    doi: 10.1017/S0033291707001638

Hemsley, D. R., & Garety, P. A. (1986). The formation of maintenance of delusions: a

    Bayesian analysis. *The British Journal of Psychiatry*, *149*, 51–56.

    https://doi.org/10.1192/bjp.149.1.51

Juárez-Ramos, V., Rubio, J. L., Delpero, C., Mioni, G., Stablum, F., & Gómez-Milán, E.

    (2014). Jumping to conclusions bias, BADE and feedback sensitivity in schizophrenia

    and schizotypy. *Consciousness and cognition*, *26*, 133-144.

    https://doi.org/10.1016/j.concog.2014.03.006

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

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical

    data. *Biometrics*, 159-174. http://www.jstor.org/stable/2529310.

Lavigne, K. M., Menon, M., Moritz, S., & Woodward, T. S. (2020). Functional brain

    networks underlying evidence integration and delusional ideation. *Schizophrenia*

    Research, *216*, 302-309. https://doi.org/10.1016/j.schres.2019.11.038

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... &

    Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-

    analyses of studies that evaluate health care interventions: explanation and

    elaboration. *Journal of clinical epidemiology*, *62*(10), e1-e34.

    https://doi.org/10.1016/j.jclinepi.2009.06.006

McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2017). Association of the jumping to

    conclusions and evidence integration biases with delusions in psychosis: a detailed

    meta-analysis. *Schizophrenia bulletin*, *43*(2), 344-354.

    https://doi.org/10.1093/schbul/sbw056

Menon, M., Quilty, L. C., Zawadzki, J. A., Woodward, T. S., Sokolowski, H. M., Boon, H.

    S., & Wong, A. H. C. (2013). The role of cognitive biases and personality variables in

    subclinical delusional ideation. *Cognitive neuropsychiatry*, *18*(3), 208-218.

    https://doi.org/10.1080/13546805.2012.692873

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moritz, S., Andreou, C., Schneider, B. C., Wittekind, C. E., Menon, M., Balzan, R. P., & Woodward, T. S. (2014). Sowing the seeds of doubt: a narrative review on metacognitive training in schizophrenia. *Clinical psychology review*, *34*(4), 358-366. https://doi.org/10.1016/j.cpr.2014.04.004

Moritz, S., Veckenstedt, R., Hottenrott, B., Woodward, T. S., Randjbar, S., & Lincoln, T. M. (2010). Different sides of the same coin? Intercorrelations of cognitive biases in schizophrenia. *Cognitive Neuropsychiatry*, *15*(4), 406-421. https://doi.org/10.1080/13546800903399993

Moritz, S., & Woodward, T. S. (2006). A generalized bias against disconfirmatory evidence in schizophrenia. *Psychiatry research*, *142*(2-3), 157-165. doi: 10.1016/j.psychres.2005.08.016

Orenes, I., Navarrete, G., Beltrán, D., & Santamaría, C. (2012). Schizotypal people stick longer to their first choices. *Psychiatry Research*, *200*(2-3), 620-628. https://doi.org/10.1016/j.psychres.2012.03.030

Pappas, C., & Williams, I. (2011). Grey literature: its emerging importance. *Journal of Hospital Librarianship*, *11*(3), 228-234. doi: 10.1080/15323269.2011.587100

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

Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on

DSM-III-R criteria. *Schizophrenia bulletin*, *17*(4), 555-564.

https://doi.org/10.1093/schbul/17.4.555

Riccaboni, R., Fresi, F., Bosia, M., Buonocore, M., Leiba, N., Smeraldi, E., & Cavallaro, R.

(2012). Patterns of evidence integration in schizophrenia and delusion. *Psychiatry*

Research, *200*(2-3), 108-114. https://doi.org/10.1016/j.psychres.2012.04.005

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built

clinical question: a key to evidence-based decisions. *Acp j club*, *123*(3), A12-A13.

PMID: 7582737

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

University Press.

Sanford, N., Veckenstedt, R., Moritz, S., Balzan, R. P., & Woodward, T. S. (2014). Impaired

integration of disambiguating evidence in delusional schizophrenia

patients. *Psychological Medicine*, *44*(13), 2729-2738.

doi:10.1017/S0033291714000397

So, S. H., Freeman, D., Dunn, G., Kapur, S., Kuipers, E., Bebbington, P., . . . Garety, P. A.

(2012). Jumping to conclusions, a lack of belief flexibility and delusional conviction

in psychosis: A longitudinal investigation of the structure, frequency, and relatedness

of reasoning biases. *Journal of Abnormal Psychology, 121*(1), 129-139.

http://dx.doi.org/10.1037/a0025297

Speechley, W. J., Ngan, E. T. C., Moritz, S., & Woodward, T. S. (2012). Impaired evidence

integration and delusions in schizophrenia. *Journal of Experimental*

Psychopathology, *3*(4), 688-701. doi: 10.5127/jep.018411

Thomas B.H., Ciliska D., Dobbins M., Micucci S. (2004). A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evidence Based Nursing*, 1(3), 176-184. doi:10.1111/j.1524-475X.2004.04006.x

Veckenstedt, R., Randjbar, S., Vitzthum, F., Hottenrott, B., Woodward, T. S., & Moritz, S. (2011). Incorrigibility, jumping to conclusions, and decision threshold in schizophrenia. *Cognitive neuropsychiatry*, *16*(2), 174-192. https://doi.org/10.1080/13546805.2010.536084

Ward, T., & Garety, P. A. (2019). Fast and slow thinking in distressing delusions: A review of the literature and implications for targeted therapy. *Schizophrenia Research*, *203*, 80-87. https://doi.org/10.1016/j.schres.2017.08.045

Wessely, S., Buchanan, A., Reed, A., Cutting, J., Everitt, B., Garety, P., & Taylor, P. J. (1993). Acting on delusions. I: Prevalence. *The British Journal of Psychiatry*, *163*(1), 69-76. doi:10.1192/bjp.163.1.69

Woodward, T. S., Moritz, S., & Chen, E. Y.H. (2006). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions: A study in an Asian sample with first episode schizophrenia spectrum disorders. *Schizophrenia Research, 83*(2-3), 297–298. https://doi.org/10.1016/j.schres.2006.01.015

Woodward, T. S., Buchy, L., Moritz, S., & Liotti, M. (2007). A bias against disconfirmatory evidence is associated with delusion proneness in a nonclinical sample. *Schizophrenia bulletin*, *33*(4), 1023-1028. https://doi.org/10.1093/schbul/sbm013

Zawadzki, J. A., Woodward, T. S., Sokolowski, H. M., Boon, H. S., Wong, A. H. C., & Menon, M. (2012). Cognitive factors associated with subclinical delusional ideation in the general population. *Psychiatry research*, *197*(3), 345-349. https://doi.org/10.1016/j.psychres.2012.01.004

Appendices

Appendix A. PICO Search Framework

	Search Criteria	Search Terms Used
Participant	Clinical or subclinical paranoid individuals	'Paranoi*' OR 'Persecut*' OR 'Delusion*' OR 'Suspici*'
	AND	
Intervention	N/A	N/A
	AND	
Control	Clinical (high) or subclinical (low) paranoia levels	'Paranoi*' OR 'Persecut*' OR 'Delusion*' OR 'Suspici*'
	AND	
Outcome	Belief flexibility measure	"Belief Flexibility" OR "Belief Updating" OR "Belief Formation" OR "Bias Against Disconfirmatory Evidence" OR "BADE" OR "Evidence Integration" OR "Maudsley Assessment of Delusions Schedule" Or "MADS" OR "Explanation of Experiences Assessment" OR "EoE"

Appendix B. Quality Assessment Tool – EPHPP

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Section Two: Empirical Study

An Experimental Study Investigating Trustworthiness Judgements in Non-Clinical Paranoia

Abstract

Objectives

Research into trustworthiness and paranoia has yielded inconsistent findings and lacks ecological validity. This study aimed to identify whether there was a difference in estimations of trustworthiness judgements, in the light of new information, for adults experiencing low or high paranoia.

Methods

A mixed between-within groups design was employed. Participants in the low and high paranoia groups completed an online scenarios-based task which included trustworthy and untrustworthy conditions. Trust judgments were assessed before and following a trust violation. Validated measures of paranoia, attachment, self-esteem, analytical thinking, and belief updating were also completed. Data were analysed using repeated measures ANOVAs

Results

Trustworthiness judgements in both conditions between the low and high paranoia group were non-significant before and following a trust violation. However, participants changed their trust judgements in response to new information in the way that was expected. In both conditions, a significant interaction between gender of faces and time, with greater trustworthiness rating for females across time, and a significant interaction between age and group, with greater paranoia severity demonstrated in younger people, was found. However, the association between high paranoia and attachment insecurity and self-esteem was non-significant.

Conclusions

The findings indicated a non-significant difference in the rate a person adjusts their estimates of trustworthiness judgements between the groups. However, the study paradigm was found to be sensitive to changing estimates of trustworthiness. Future research may benefit from using this paradigm with clinical and non-clinical samples to further the understanding of paranoia and trustworthiness.

Keywords: paranoia, persecutory, trustworthiness, trust judgments

Practitioner points:

- Estimates of trustworthiness judgements do not appear to differ between the low and high paranoia groups.
- The study paradigm is sensitive to changing estimates of trustworthiness, but further research is required with clinical and non-clinical samples.
- Practitioners are likely to improve their ability to form therapeutic alliances if they pay attention to the way that trust evolves within the therapeutic sessions.

Introduction

Paranoia

Paranoia can be defined as unfounded or exaggerated beliefs relating to other people and to the world which is characterised by suspicion, mistrust, or hostility (Freeman et al., 2005). Paranoia is the most common symptom of psychosis (Bentall et al., 2001). However, less severe levels can be found within a non-clinical population, with 10-15% of the general population estimated to experience some degree of paranoia (Freeman et al., 2005).

Paranoia can be understood as a continuum and described using the paranoia hierarchy model (Freeman et al., 2005; Figure 1). Social evaluative concerns such as fear of rejection, which are commonly experienced, are bottom of the hierarchy, whilst severe threats, for example, the perception that others are causing the individual significant harm, which is less common, are top of the hierarchy. This model is supported by psychometric research which has evidenced the hypothesis of a continuum running from ordinary suspiciousness through to paranoid delusions (Elahi et al., 2017).

Figure 1

Hierarchy of Paranoia
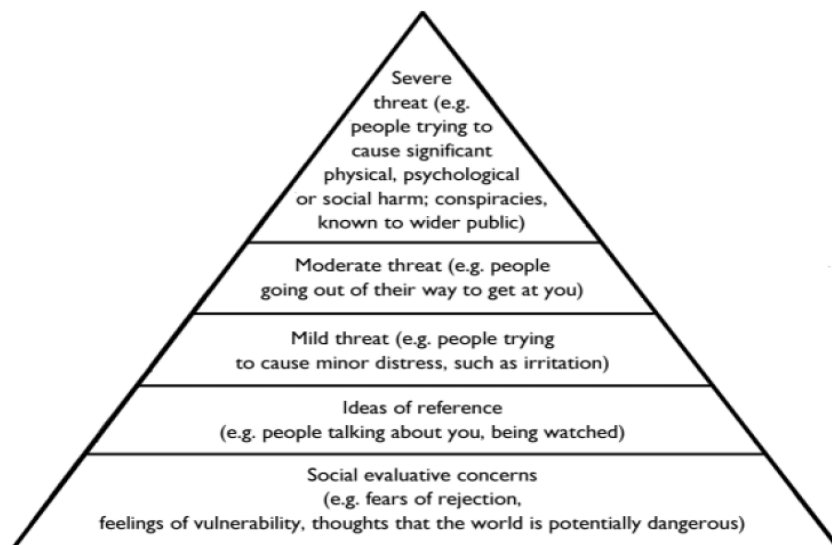
Attachment, Self-Esteem, and Paranoia

There is growing evidence that attachment style and low self-esteem are strong predictors of paranoia (Lavin et al., 2020; Thewissen et al., 2011). Experiences of childhood adversity in early years are associated with attachment insecurity (Bowlby, 1969; Kim et al., 2021; Sitko et al., 2014). Accordingly, specific causal pathways from early insecure attachment and the experience of paranoia have been proposed (Bentall et al., 2014; Bentall & Fernyhough, 2008; Lavin et al., 2020; MacBeth et al., 2011).

Research has consistently shown an association between paranoia and insecure attachment. Wichham et al. (2015) found that within a clinical sample, attachment avoidance and attachment anxiety were both associated with paranoia. Similarly, within the public, attachment insecurity predicted paranoia and this relationship remained robust after controlling for comorbidity (Pickering et al., 2008). Consequently, suggesting an insecure attachment style developed in early childhood is a key contributor to the experience of paranoia later in life.

A recent meta-analysis of twenty-six studies conducted by Murphey et al. (2020) found a significant association between paranoia and both attachment anxiety and attachment avoidance, with similar strength associations for clinical and non-clinical samples. The review concluded that attachment insecurity is a contributing factor to the development and/or maintenance of paranoia.

Research has also found an association between paranoia and self-esteem. Lower self-esteem was a prominent feature of paranoia in clinical samples (Freeman et al., 1998), non-clinical samples (Combs & Penn, 2004; Ellett et al., 2003; Martin & Penn, 2001), and found to play a mediating role between insecure attachment and paranoia in both non-clinical and clinical samples (Pickering et al., 2008; Ringer et al., 2014; Wickham et al., 2015).

A recent systematic review of forty-three studies and a meta-analysis of twenty-five studies examined the relationship between negative self- and other schemas in paranoia (Humphrey et al., 2021). The findings indicated a medium to large relationship between the variables, suggesting negative self- and other schemas correlate with paranoia. Overall, the findings provide support for the association between self-esteem and self-/other schemas and paranoia.

Given the robust association between paranoia and both insecure attachment style and low self-esteem, measures of both constructs were included in the study as a test of the validity of the method used to assign participants to low and high paranoia groups.

Jumping to Conclusions, Analytical Reasoning, and Paranoia

Jumping to conclusions (JTC) is a style of cognitive reasoning whereby decisions are made on limited information, without consideration of alternatives or gaining further information (Freeman et al., 2014a). JTC is often measured using a probabilistic reasoning task, for example, the beads task. Two jars predefined with an equal but opposite ratio of coloured beads (85 blue and 15 red beads or vice versa) are presented. Beads are selected sequentially from a jar and participants guess the jar the beads have been drawn from. Guesses made within two or fewer beads are considered JTC (Garety et al., 2005)

A growing literature suggests people experiencing delusional ideation are more likely to have JTC compared to controls (Freeman 2008; Fine et al., 2007). However, non-clinical samples experiencing delusions have been seldom studied with inconclusive findings (Colbert & Peters, 2002; McKay et al., 2006; Van Dael et al., 2006).

Fewer studies have investigated subtypes of delusions or paranoia on JTC. Startup et al. (2008) found a correlation between JTC and persecutory delusions. Similar findings were noted by Garety et al. (2013) but a stronger association was reported for grandiose delusions

compared to persecutory delusions with an 85:15% variant for beads. Additionally, Bentall et al. (2009) found that JTC was associated with paranoia but so were other measures of executive functioning, therefore, cognitive impairment in paranoia was not restricted to the specific domain measured by the JTC task.

Freeman and colleagues (2012) found that increased intuitive thinking and reduced reflective thinking were associated with persecutory ideation, however, the associations were small. Similarly, lower analytic reasoning was found in subclinical paranoia (Freeman et al., 2014b) and people experiencing delusions (Speechley et al., 2010; Speechley et al., 2013). Consequently, JTC may predict delusion persistence and maintenance (Dudley et al., 2013; Menon et al., 2008).

Given the association between paranoia and both JTC and analytical reasoning, measures of both constructs were included in the study as a test of the validity of the method used to assign participants to low and high paranoia groups.

Trustworthiness

Trust plays an important role in social judgements, interactions, and economic well-being (Rempel et al., 1985; Zak & Knack, 2001). The Oxford English Dictionary (1989) defines *Trust* as, "Confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement." This definition has become synonymous with the term Trustworthiness.

Trustworthiness evaluations are based on three characteristics: 1) Competency, assessment of an individual's ability to manage complex situations; 2) Integrity, consideration of an individual's moral principles and honesty, and 3) Intentions, an individual's goodwill or intent (Colquitt et al., 2007; Mayer, et al., 1995; Rousseau et al., 1998). These trustworthiness evaluations are based on prior knowledge of an individual.

However, people quickly assess trustworthiness during a first encounter, therefore before the aforementioned information has been acquired (Willis & Todorov, 2006; Winston et al., 2002;).

Research suggests people use faces to make trustworthiness judgments as facial expressions provide a wealth of social information (Adolphs, 2002). Interestingly, these judgements, which are made at the rate of 100 milliseconds following exposure (Willis & Todorov, 2006), are influenced by the gender of the face. Female faces are considered more trustworthy than male faces; with features stereotypically associated with females rated higher for trust (Sutherland et al., 2015; Wincenciak et al., 2013).

However, neutral gendered faces can be interpreted in several ways, for example, as trustworthy and attractive (Oosterhof & Todorov, 2008). Faces that are perceived as either intelligent, attractive, or non-aggressive tend to be regarded as trustworthy (Willis & Todorov, 2006).

Trustworthiness and Paranoia

Trustworthiness judgements are important in day-to-day interactions especially as they support decision making, for example, whether to approach someone new (Hajdúk et al., 2019). However, the ability to formulate trustworthiness judgements differs between individuals. People who experience paranoia have been found to make less accurate judgements of people they meet for the first time (Kirk et al., 2013).

Kirk et al. (2013) investigated trustworthy evaluations of unfamiliar faces manipulated to look untrustworthy, neutral, and trustworthy, in a non-clinical sample of low and high paranoia. Despite both groups changing their rating in accordance with the trustworthiness of faces, the high paranoia group rated unfamiliar faces significantly less trustworthy than the low paranoia group. It is believed that individuals experiencing paranoia

have an interpretation bias towards threat or perceived danger, which occurs even when there is limited supporting evidence. Accordingly, individuals experiencing paranoia have a preferential recall of threat-related information (Kaney et al., 1992; Phillips et al., 2000; Trotta et al., 2020).

However, some studies produced varying results when using both static judgments of faces and dynamic money exchange tasks. One study investigating trustworthiness in patients diagnosed with schizophrenia found unfamiliar faces were rated more trustworthy in comparison to the control group (Baas et al., 2008). This difference was explained by the idea of reduced social cognitive abilities in participants with schizophrenia (Green et al., 2008a).

Conversely, in a more recent study, participants with high paranoia ideation rated computer-generated faces, previously calibrated as either trustworthy or untrustworthy, to be less trustworthy compared to the low paranoia group. It would therefore suggest a bias towards mistrust in people experiencing greater paranoia severity (Martinez et al., 2021).

Furthermore, Reed et al. (2020) investigated belief updating in paranoia. The study found that increased paranoia maintained initial biases even when provided with contrary evidence. This would suggest that individuals experiencing increased paranoia severity have difficulty updating beliefs which could be explained by their bias towards mistrust (Kirk et al., 2013).

Current Study Rationale

Lack of trust is considered a component of paranoia, but it is not sufficient to define paranoia, which also includes interpersonal vulnerability and suspiciousness (Bebbington, 2013). Research on healthy individuals shows that trust is a dynamic process which changes with time and experiences (Willis & Todorov, 2006). Previous studies have used static perceptual judgement tasks or dynamic monetary exchange tasks to understand interpersonal

trustworthiness and paranoia. However, the approaches lack ecological validity of everyday interactions. The current study, therefore, provides a novel and dynamic approach, using a specific scenarios-based task, to understand the way that people with low and high levels of paranoia change their trustworthiness judgements as new information becomes available.

Aim

This study aims to identify whether there is a difference in the rate a person experiencing high paranoia adjusts their estimations of trustworthiness in the light of new information compared to a person experiencing low paranoia.

Hypothesis

1. People experiencing high paranoia will be quicker to learn that neutral faces are untrustworthy due to their previously demonstrated bias.

2. People experiencing high paranoia will be slower to learn that neutral faces are trustworthy due to their previously demonstrated mistrust bias.

3. When encountering a trust violation (somebody who has behaved in a trustworthy fashion suddenly behaves in an untrustworthy fashion), people experiencing high paranoia will show greater reductions in trustworthiness judgements compared to previous estimates whereas, in the low paranoia group, these judgements will be less affected.

4. When encountering a mistrust violation (somebody who has behaved in an untrustworthy fashion suddenly behaves in a trustworthy fashion), people experiencing high paranoia will show smaller increases in trustworthiness judgments compared to the low paranoia group.

Secondary Hypothesis

5. People in the high paranoia group will present with attachment insecurity and low levels of self-esteem and these will predict changes in trustworthiness judgements observed in the two groups.

Methods

Design

The current study employed a quasi-experimental between-within-groups design. This approach allowed the examination of the differences in the rate of change of trustworthiness judgements between the low and high paranoia groups as well as within the low and high paranoia groups.

Participants

Power Analysis

A statistical power analysis was performed for sample size estimation based on data from a pilot study ($n = 90$) that measured bias towards mistrust using a repeated measure design (Martinez et al., 2021). The effect size in the aforementioned study for bias towards mistrust was small, $d = .03$ (Cohen, 1998).

With an alpha of .05 and power at .80, the projected sample size needed with this effect size (GPower 3.1) was approximately 56 participants for a simple between/within-group comparison. Thus, the proposed sample size of 120 was deemed more than adequate for the main objective of the study.

Recruitment

The present study recruited a sample of 490 online British participants. All participants were screened using the GPTS-B (Green et al., 2008b; Appendix A). A cut-off

total score of equal to or greater than 35 was used to assign people to the high paranoia

group, whilst a cut-off total score of equal to or less than 34 was used to assign people to the

low paranoia group, as recommended by Freeman et al. (2019). The recruitment of the low

and high paranoia participants adopted different strategies, via social media ($n = 115$) and

Prolific ($n = 385$), respectively.

Low and High Paranoia Recruitment

The low paranoia group was recruited via opportunity sampling through social media

platforms, including Facebook and Instagram, and participants had the option to be placed in

a prize draw for a £25 Amazon voucher. Social media platforms enabled greater

dissemination of the study and timely recruitment of the target population. During this

recruitment, if participants with high paranoia were identified, they were assigned to the high

paranoia group and their results were included in the study.

The high paranoia group was recruited via opportunity sampling through Prolific and

participants were rewarded for their participation (£0.42 for screening and £3.75 for the main

study). Prolific was used due to the difficulties recruiting the target population given the

evidence that non-clinical, high paranoia participants typically represent about 10-15% of the

population (Freeman et al., 2019).

Inclusion and Exclusion Criteria

The current study used predefined inclusion and exclusion criteria. Participants who

were adults, aged 18 and over, who could read and understand English were included.

Excluded participants were those who did not consent to partake, had missing demographics

or did not rate trustworthiness in the computerised trials (Appendix B).

Materials

Paranoid Thought Scale

The Green's Paranoid Thought Scale-B (GPTS-B; Green et al., 2008b; Appendix A) is a trait measure of paranoia. The paranoia subscale-B was used in the current study. The scale consists of sixteen items, for example, "People have intended me harm", which are assessed over the last month. The items are rated on a 5-point Likert scale, ranging from 1-Not at all and 5-Totally, which correspond to their thoughts of others. Scores for the sub-scale range from 16-80 with higher scores indicating greater paranoia ideation.

The GPTS-B has good psychometric properties within clinical and non-clinical populations. Moreover, the subscale shows excellent internal consistency ($\alpha = .92$; within the current study $\alpha = .96$.) and good test-retest reliability (Green et al., 2008b; Bronstein et al., 2019). Furthermore, the validity of the GPTS-B is maintained within a non-clinical sample as higher GPTS-B scores were associated with an increased likelihood of reporting paranoia during a virtual reality study (Freeman et al., 2010).

Self-Esteem Rating Scale–Short Form

The Self-Esteem Rating Scale-Short Form (SERS; Lecomte et al., 2006; Appendix C) is a 20-item self-report measure assessing self-esteem irrespective of mood. The scale consists of 10 positive statements, for example, "I feel I am a very competent person", and 10 negative statements, for example, "I feel inferior to other people", about the self. The items are rated on a 7-point Likert scale ranging from 1-Never to 7-Always, with higher scores representing higher self-esteem. The scale reliability is good for both positive self-esteem ($\alpha = .94$; current study $\alpha = .93$) and negative self-esteem ($\alpha = .94$; current study $\alpha = .93$).

The rationale to use this measure stemmed from the literature evidencing negative self-esteem as a strong predictor of paranoia (Thewissen et al., 2011). Consequently, the measure outcomes will provide additional support for the validity of the current study.

Relationship Questionnaire

The Relationship Questionnaire (RQ; Bartholomew & Horowitz, 1991; Appendix D) assesses adult attachment styles based on the four-category framework (secure, fearful-avoidant, preoccupied, and dismissing-avoidant; Bartholomew, 1990). Individuals are requested to identify which attachment style best describes them. The questionnaire was used to calculate the two underlying dimensions, attachment anxiety and attachment avoidance (Griffin & Bartholomew, 1994). Positive scores on attachment representations indicated more positive models whilst negative scores indicated more negative models. The RQ has high reliability and validity ranging from $\alpha = .87\text{-}.95$ for each attachment orientation (Bartholomew and Horowitz, 1991).

The rationale to use this measure was based on the evidence that attachment style is significantly associated with paranoia (Lavine et al., 2020). Consequently, the measure outcomes will provide further evidence to support the validity of the current study.

Cognitive Reflection Test

The Cognitive Reflection Test (CRT; Frederick, 2005; Appendix E) measures analytical thinking using math problems. For example, "A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost?". Each question has four response options including an intuitive, incorrect response (£0.10) and a reflective, correct response (£0.05). Scores are calculated by the number of correct responses, with higher scores reflecting greater cognitive ability.

The current study employed a five-item CRT formed of the original items ($n = 3$; Frederick, 2005) and CRT-2 items ($n = 2$; Thomson & Oppenheimer, 2016) and showed good reliability $\alpha = .77$. The two problems from the CRT-2 were included as they both yielded lower variability in response options and calculation ability, suggesting a stronger probability of intuitive or reflective responses (Thomson & Oppenheimer, 2016). An additional question "Have you seen these brain puzzles before?" was included to account for the familiarity of the puzzles.

The rationale for using the CRT was based on findings reporting intuitive thinking rather than reflective/deliberate thinking was associated with greater unusual ideation, therefore, providing support to the validity and reliability of the computerised task (Freeman et al., 2012; Ward & Garety, 2017; Ward et al., 2018).

Jumping to Conclusions Task

Jumping to conclusions bias (JTC) was assessed using an adapted computerised probabilistic reasoning "beads task" (Phillips & Edward, 1966; Appendix F). Participants were presented with two bead jars, a mostly orange jar, Jar A, (60% Orange:40% Purple) and a mostly purple jar, Jar B, (40% Orange:60% Purple). Both jars remained on the screen and a bead was drawn. The task was to identify the jar the beads were drawn from. The predetermined sequence of the beads ($n = 15$) followed the order: P-P-P-O-P-P-O-O-O-P-P-P-P-P-P. There was also a reference point at the top of the page indicating the beads drawn. After each draw, participants could "see another bead" or "guess the jar". The variable of interest was the number of beads drawn before guessing the jar. JTC was identified when a decision was made within two or fewer beads (Garety et al., 2005).

The rationale for including the JTC task was based on a meta-analysis which found that in comparison to controls, people with delusions requested significantly fewer beads

before guessing the jar (Ross et al., 2015). Consequently, the validity and reliability of the computerised task within the current sample could be established.

Facial Trust Detection Task

The Karolinska Directed Emotional Faces (KDEF; Lundqvist et al., 1998) dataset was used to obtain the faces for the scenario-based task. Four randomly selected neutral faces, two male and two female faces were used (Appendix G). Each male and female face was assigned to a predefined trustworthy and untrustworthy condition. Male 1 and Female 1 were assigned to condition 1 and condition 3, respectively, the trustworthy conditions. Male 2 and Female 2 were assigned to condition 2 and condition 4, respectively, the untrustworthy conditions. To account for order-effects participants were shown one of the four faces at random during the computerised task. All faces were presented at a 0-degree and forward-facing. The decision to use neutral faces ensured participant bias in facial judgment did not confound the results.

Scenarios

Six scenarios were used for each of the four faces, resulting in 24 scenarios, for example, "You are standing next to this person who is also waiting for the train. You realise you have left your bank card in the ticket office". Each scenario also had a scenario-specific question, for example, "How trustworthy would you say this person is to watch your luggage whilst you collect your bank card?" (Appendix H).

It was perceived that six scenarios provided enough opportunity for the participants to learn the level of trustworthiness. The scenarios, from trials 1-6 are based within three trustworthiness blocks, low (trial 1-2, for example, "How trustworthy would you say this person is to watch your luggage whilst you collect your bank card?"), medium (trial 3-4, for example, "How trustworthy would you say this person is with your details to help you access the airport Wi-Fi?"), and high (trial 5-6, for example, "How trustworthy would you say this

person is to wait for you with your passport and prevent the gates from closing"). As the trials progressed for each condition, the level of trust required from the participant also increased. This was to understand how the participant's initial learning (from previous scenarios within the condition) subsequently determined their trust judgement when the scenarios demanded greater trust.

A violation scenario was also added at trial 5 for both conditions (Appendix H). In the trustworthy conditions, both the male and female characters presented as trustworthy until trial 5 and then suddenly became untrustworthy, therefore leading to a negative trust violation and vice versa for the untrustworthy condition, leading to a positive trust violation.

A violation scenario was added based on the understanding that belief updating may be biased by previous estimates of trustworthiness for people with high paranoia. Consequently, the violation scenarios assessed whether a change in trustworthiness ratings occurred between trial 5 and trial 6 for both conditions (trustworthy and untrustworthy) in both groups.

Trustworthiness Likert Scale

The trustworthiness scale was a study-specific 7-point Likert scale ranging from 1-Very Untrustworthy to 7-Very Trustworthy. Greater graduations were included as this increases the reliability and validity of the scale (Lozano et al., 2008; Weijters et al., 2010).

Each scenario, with the assigned face, scenario relevant question, and the trustworthiness rating scale remained on the screen until the participant had rated their level of trustworthiness and manually clicked next. Thereafter, the participant was provided with an outcome related to the scenario, for example, "You returned from the ticket office and see that the person has watched your luggage, and nothing is missing" (Figure 2).

Figure 2

Scenario Presentation Screenshot

Procedure

Pilot

Service user involvement was not required as the participants of interest were the public. A non-clinical sample was recruited to pilot the current study to assess the feasibility of the approach and identify any limitations (Hassan et al., 2006).

Twelve members of the public, predominantly female ($n = 7$), were recruited via Instagram to take part. The age range of the sample was 24-59 years ($M = 34.9$ $SD = 12.1$). The 'Rule of 12' participants for a pilot study is deemed appropriate in providing valuable

information for planning a larger subsequent study (Moore et al., 2011). During the pilot, the study protocol was followed closely. Recommendations, for example, the wording of the scenarios were attended to before commencing the current study (Appendix I).

Data Collection

The study was programmed on Qualtrics, an online survey platform. Prolific and Social Media platforms were used to distribute the study information including the Qualtrics link (Appendix J). Each participant was provided with a recruitment-specific information sheet outlining the purpose of the study (Appendix K and L). Demographic characteristics and consent were also obtained (Appendix M). Participants recruited via Prolific, meeting the inclusion criteria, were invited back for the main study and reminded of their consent and right to withdraw (Appendix N). The recruitment process and participation in the study ran simultaneously for nine months (June 2021-March 2022).

All participants, according to the recruitment approach (Figure 3), completed three standardised measures, GPTS-B, SERS-Short Form, and RQ, as well as two reasoning tasks CT and JTC, and the computerised task.

Before commencing the computerised task, written instructions were provided electronically outlining the process of the study (Appendix O). Participants completed baseline ratings of trustworthiness using a 7-point Likert scale, 1-Very Untrustworthy to 7-Very Trustworthy, for each face. Each face and the Likert scale remained on the screen until the participant responded.

During the computerised task, a randomised face, along with the scenario, and a scenario-specific question were presented. Participants provided trust judgment ratings on a 7-point Likert scale, 1-Very Untrustworthy to 7-Very Trustworthy, for the question presented on the screen, for example, "Would you trust this person to watch your bags?". All

information remained on the screen until the participant had responded. Following the trust rating, participants were presented with a scenario outcome, for example, "The person watched your bags, and nothing is missing".

Participants worked through 24 randomly presented scenarios ascending from low- to high-level trust for the remaining 3 faces (6 trials per face). Upon completion, participants were provided with a debrief form and thanked for their participation (Appendix P). The study duration from start to finish was approximately 30 minutes.

Figure 3

Participant Allocation and Study Procedure for the Low and High Paranoia Group
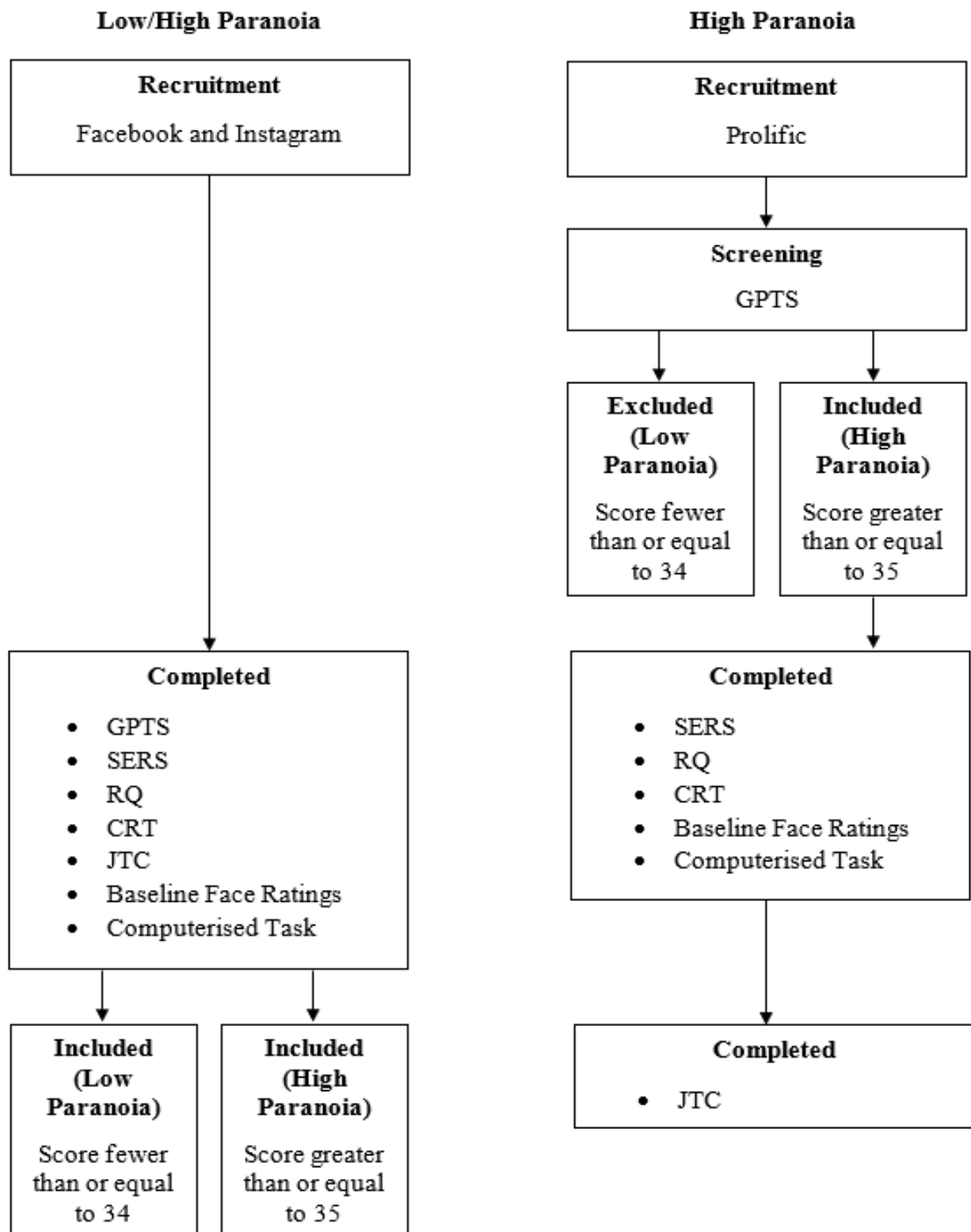
Ethics

This study was granted ethical approval by the Psychology Ethics Committee at The University of Sheffield (Appendix Q). Key ethical considerations were attended to including informed consent and the right to withdraw. A debrief form including the researcher's contact details and mental health support information was also provided (Appendix P).

Results

Data Analytic Strategy

Data were analysed using the IBM statistics 26 package. Independent samples t-test and Pearson's chi-squared test were used to verify any differences between the two groups (low and high paranoia) on demographic or clinical variables.

To simplify the analysis and facilitate the interpretation of the findings, two 3-way repeated measure analysis of variance (ANOVA) were conducted. The first ANOVA focused on the trustworthy conditions whilst the second ANOVA focused on the untrustworthy conditions. In each ANOVA there were 3 factors. These were the two groups (*low and high paranoia*), the gender of the faces (*male and female*), and the repeated measure was the scenarios on the paradigm (*Baseline-5*). A 4-way ANOVA with an additional factor condition (*trustworthy and untrustworthy*) yielded comparable results (Appendix R).

A main effect for group and an interaction between group and time in each of the conditions was expected. Although the gender of the faces did not pertain to the main hypothesis, a main effect of gender of the face was expected as female faces are perceived to be more trustworthy than male faces (Sutherland et al., 2015; Wincenciak et al., 2013). However, a significant interaction between the gender of the face and either group or time was not expected. Post hoc analysis using pairwise comparisons, using the Bonferroni

correction, was also completed to assess differences in trustworthy ratings between timepoints (baseline–trial 5) for both conditions.

To test how the groups responded to the trust violations, a 3-way ANOVA was completed for group (*low and high paranoia*), the gender of the faces (*male and female*) and time using the trust judgements at trial 6 minus trial 5 for both conditions.

Secondary analysis, using Pearson's correlation assessed the association between the variables of interest. The strength of the correlations will be reported using the *r* values proposed by Evans (1996). Additionally, regression analysis examined whether attachment insecurity and low levels of self-esteem predict changes in trustworthiness judgements observed in the two groups. To facilitate understanding, the current study will only report significant correlations at $p < .05$ or $p < .01$.

Participants

The study used a sample of 178 non-clinical individuals with paranoia (low paranoia $n = 100$ and high paranoia $n = 78$). Across the sample, participants were predominantly female $n = 105$ (59%). The mean age of the total sample was 45.2 years ($SD = 16.9$, *Range* = 18-92 years). Table 1 lists demographic characteristics and clinical variables for both groups.

On average, participants in the low paranoia group were older compared to the high paranoia group, with a highly significant difference $t(176) = -6.62$, $p < .001$. Additionally, there was a significant difference for gender between the two groups, with more females in the low paranoia group and more males in the high paranoia group $X^2 (2, N = 178) = 10.82$, $p < .005$. However, there was no significant difference for ethnicity ($p = .65$) and mental health support ($p = .23$) between the two groups.

Regarding clinical characteristics, there was a highly significant difference between the GPTS-B total score for the low and high paranoia group, $t(129.70) = -22.39$, $p < .001$. This was a difference by design.

A significant difference was also found for attachment anxiety, with lower attachment anxiety in the low paranoia group compared to the high paranoia group $t(176) = 4.30$ $p < .001$. However, attachment avoidance was not significantly different between the groups ($p = .32$).

Self-esteem was significantly higher in the low paranoia group compared to the high paranoia group, $t(176) = 5.10$, $p < .001$. However, the difference between the two groups for CRT was non-significant ($p = .77$).

A total of $n = 165$ participants (low paranoia $n = 97$ and high paranoia $n = 68$) completed the JTC task. Missing data was evident for the low paranoia ($n = 3$) whilst the high paranoia group experienced dropout ($n = 10$). Dropout was due to the JTC task being missed from the survey and completed after the main study. There was a significant difference for draws-to-decision, with greater draws-to-decision in the low paranoia group compared to the high paranoia group, $t(161.01) = 2.34$, $p < .05$. Nevertheless, there was a non-significant difference in the choice of jar (A or B) between the groups ($p = .29$).

Table 1

Demographic and Clinical Variables for the Low and High Paranoia Group.

Variables	Low Paranoia (*n* = 100)	High Paranoia (*n* = 78)
Demographic Variables	*n* (%)	*n* (%)
Age (years)		
Mean *(SD)*	51.85 *(15.50)*	36.73 *(14.63)*
Range	19-92	18-82
Gender		
Male	30 (30%)	42 (53.8%)
Female	69 (69 %)	36 (46.2%)
Non-Binary	1 (1%)	0
Ethnicity		
White British/Irish	60 (60%)	58 (74.4%)
White Non-British/Irish	6 (6%)	5 (6.4%)
Chinese	2 (2%)	1 (1.3%)
Indian	24 (24%)	6 (7.7%)
Pakistani	0 (0%)	2 (2.6%)
African	1 (1%)	2 (2.6%)
Other Ethnicity		
Anglo Indian	1 (1%)	0
Black African White European	0	1 (1.3%)
British Indian	0	1 (1.3%)
Irish	1 (1%)	0
Mixed	1 (1%)	0
Mixed Asian and White	1 (1%)	0
Mixed White and Black Caribbean	1 (1%)	2 (2.6%)
Mixed White Non-British/Bangladeshi	1 (1%)	0

Variables	Low Paranoia	High Paranoia
	(*n* = 100)	(*n* = 78)
Mental Health Support	*n* (%)	*n* (%)
Never received	55 (55%)	35 (44.9%)
Previously received	32 (32%)	26 (33.3%)
Currently receiving	12 (12%)	17 (21.8%)
Prefer not to say	1 (1%)	0
Clinical Variables		
GPTS-B Total		
Mean *(SD)*	20.19 *(6.40)*	47.97 *(9.39)*
Range	15-34	35-73
RQ Self-Anxiety		
Mean *(SD)*	1.75 *(4.33)*	-1.06 *(4.34)*
RQ Other Avoidance		
Mean *(SD)*	.25 *(4.04)*	-1.58 *(4.24)*
SERS-Short Form Total		
Mean *(SD)*	95.57 *(19.31)*	79.86 *(21.72)*
Range	51-138	37-122
CRT Total		
Mean *(SD)*	2.68 *(1.73)*	2.60 *(1.73)*
CRT Seen Before		
None *n (%)*	54 (54%)	29 (37.2%)
Some *n (%)*	46 (46%)	49 (62.8%)
JTC Draws to Decision	*n* = 97	*n* = 68
Mean *(SD)*	3.28 (3.33)	2.29 (2.08)

Note. CRT = Cognitive Reflection Test, GPTS-B = Green's Paranoia Thought Scale –B, JTC = Jumping to Conclusions, RQ = Relationship Questionnaire, SERS-Short Form = Self-Esteem Rating Scale- Short Form.

Baseline Scores

The average baseline rating scores for both the male and female faces in the trustworthy and untrustworthy condition were high (more trustworthy) for the low paranoia group compared to the high paranoia group (Table 2). Although the means were higher in the low paranoia group, the multivariate analysis of variance did not find a significant main effect of group ($p = .09$).

Table 2

Baseline Face Ratings for the Low and High Paranoia Group.

Condition	Faces	M	SD
Trustworthy	Male		
	Low Paranoia	4.00	1.17
	High Paranoia	3.51	1.33
	Female		
	Low Paranoia	4.32	1.21
	High Paranoia	4.08	1.34
Untrustworthy	Male		
	Low Paranoia	4.18	1.20
	High Paranoia	3.90	1.49
	Female		
	Low Paranoia	4.25	1.22
	High Paranoia	4.22	1.31

Hypothesis 1 and 2

Trustworthy Condition

Mauchly's test of sphericity was violated for time (X^2 (14) = 139.15, $p < .01$) and the time and gender of face interaction (X^2 (14) = 136.76, $p < .01$) for the trustworthy condition. The Greenhouse-Geisser estimate of sphericity was therefore used to correct the degrees of freedom ($\varepsilon = .75$ and .89, respectively).

The repeated measures ANOVA revealed a highly significant main effect of time $F(3.72, 655.22) = 107.01$, $p < .001$, $\eta_p^2 = .38$ and gender of face $F(1,176) = 36.25$, $p < .001$, $\eta_p^2 = .17$. The time and gender of face interaction was also highly significant $F(4.46, 784.96) = 11.04$, $p < .001$, $\eta_p^2 = .06$. The 3-way ANOVA for the trustworthy condition was non-significant $F(5,172) = 2.19$, $p = .058$, $\eta_p^2 = .06$. All remaining main effects and interactions were non-significant ($p > .05$). Post hoc analysis found all timepoints (baseline-trial 5) were significant (all p's $< .05$) except for timepoints 1-2 ($p = 1.00$), 4-5 ($p = 1.00$), and 4-6 ($p = .11$) which were non-significant.

Untrustworthy Condition

Mauchly's test of sphericity was violated for time (X^2 (14) = 165.16, $p < .01$) and the time and gender of face interaction (X^2 (14) = 60.54, $p < .01$). The Greenhouse-Geisser estimate of sphericity was therefore used to correct the degrees of freedom ($\varepsilon = .74$ and .87, respectively).

The repeated measures ANOVA revealed a highly significant main effect of time $F(3.68, 647.60) = 109.87$, $p < .001$, $\eta_p^2 = .38$ and gender of face $F(1,176) = 55.91$, $p < .001$, $\eta_p^2 = .24$. The time and gender of face interaction was also significant $F(4.37, 768.23) = 2.40$, $p < .05$, $\eta_p^2 = .013$. The 3-way ANOVA for the untrustworthy condition as well as the remaining interactions were all non-significant ($p > .05$). Post hoc analysis found all

timepoints (baseline-trial 5) to be significant (all p's $< .05$) except for timepoints 1-3 ($p = .20$) and 4-6 ($p = 1.00$) which were non-significant.

The significant main effect of gender of face for both analyses was accounted for by the female faces rated more trustworthy than the male faces. The significant effect of time for both conditions indicates that the scenarios from baseline to trial 5 were rated differently, with higher ratings of trustworthiness or untrustworthiness in the trustworthy and untrustworthy conditions, respectively (Table 3; Figures 4 and 5, respectively). The significant interaction between gender of face and time suggests males and females were rated differently across time for both conditions with a greater trustworthiness rating for females across time.

Table 3

Baseline to Trial 5 Mean Trustworthiness Ratings for the Trustworthy and Untrustworthy Conditions

Time	Trustworthy Condition								Untrustworthy Condition							
	Low Paranoia				High Paranoia				Low Paranoia				High Paranoia			
	Male		Female		Male		Female		Male		Female		Male		Female	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Baseline	4.00	1.17	4.32	1.21	3.51	1.33	4.08	1.34	4.18	1.20	4.25	1.22	3.90	1.49	4.22	1.31
1	3.59	1.38	4.75	1.38	3.46	1.61	4.47	1.27	4.52	1.36	5.10	1.19	4.53	1.46	5.10	1.32
2	4.51	1.41	4.99	1.35	4.00	1.71	4.72	1.49	3.60	1.62	4.35	1.45	3.41	1.72	4.05	1.62
3	5.45	1.27	5.59	1.11	5.50	1.31	5.28	1.40	2.52	1.47	3.23	1.54	2.77	1.87	3.09	1.68
4	5.40	1.42	5.79	1.18	5.54	1.51	5.40	1.68	2.38	1.50	2.94	1.73	2.21	1.53	2.92	1.85
5	4.85	1.47	5.46	1.41	5.06	1.60	5.50	1.44	2.89	1.53	2.94	1.64	2.82	1.88	3.44	2.07

Figure 4

Trust Ratings from Baseline to Trial 5 in the Trustworthy Condition
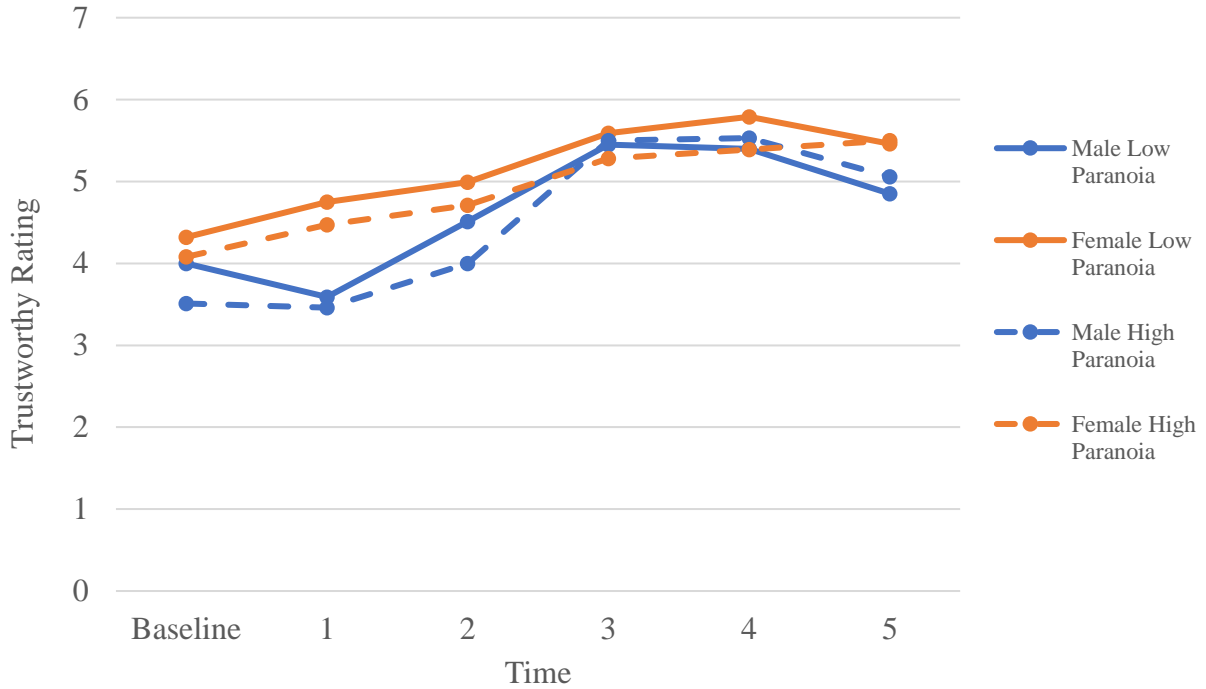
Figure 5

Trust Ratings from Baseline to Trial 5 in the Untrustworthy Condition
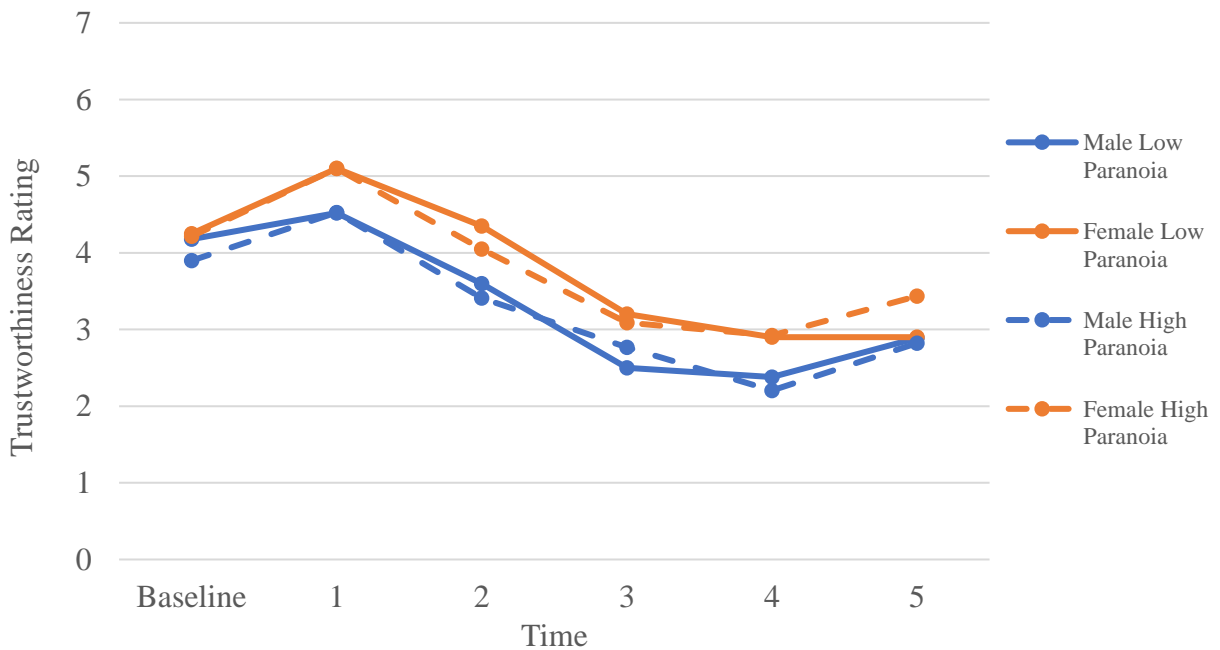
Hypothesis 3 and 4

Trustworthy Condition

The violation from trustworthy to untrustworthy from trial 5 to trial 6 was assessed for both the male and female faces in the trustworthy conditions. The average scores for both faces on trials 5 and 6 are shown in Table 4 and Figure 6. A 3-way ANOVA revealed a significant effect for gender of face $F(1,176) = 117.60$, $p < .001$, and a significant effect for time $F(1,176) = 32.40$, $p < .001$. However, there was no significant interaction between group, gender of face, and time ($p = .93$).

Untrustworthy Condition

The violation from untrustworthy to trustworthy from trial 5 to trial 6 was assessed for both the male and female faces in the untrustworthy conditions. The average scores for both faces on trials 5 and 6 are shown in Table 4 and Figure 7. A 3-way ANOVA revealed a significant effect for gender of face $F(1.176) = 5.88$, $p < .05$, and a significant effect for time $F(1,176) = 6.26$, $p < .05$. However, there was no significant interaction between group, gender of face, and time ($p = .15$).

Table 4

Average Ratings for Trial 5 and Trial 6

Condition	Faces	Trial 5 M (SD)	Trial 6 M (SD)
Trustworthy	Male		
	Low Paranoia	4.85 *(1.47)*	3.71 *(1.61)*
	High Paranoia	5.06 *(1.60)*	3.63 *(1.95)*
	Female		
	Low Paranoia	5.46 *(1.41)*	4.51 *(1.45)*
	High Paranoia	5.50 *(1.44)*	4.22 *(1.70)*
Untrustworthy	Male		
	Low Paranoia	2.89 *(1.53)*	3.20 *(1.45)*
	High Paranoia	2.82 *(1.88)*	3.15 *(1.74)*
	Female		
	Low Paranoia	2.94 *(1.64)*	3.44 *(1.41)*
	High Paranoia	3.44 *(2.07)*	3.38 *(1.51)*

Figure 6

Trust Ratings from Trial 5 to Trial 6 in the Trustworthy Condition
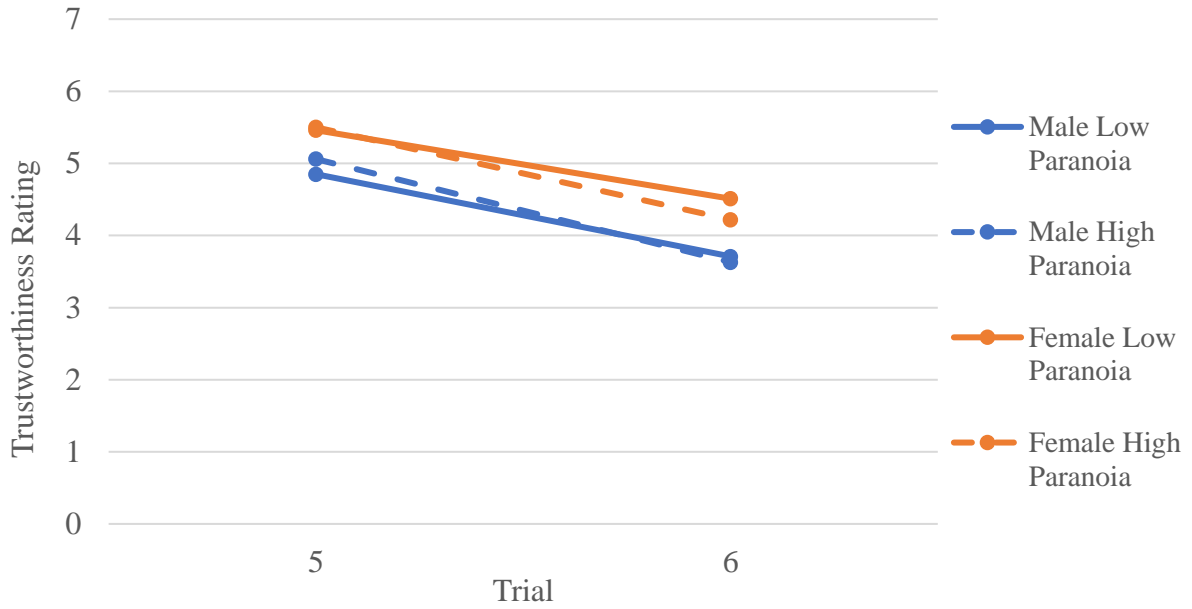
Figure 7

Trust Ratings from Trial 5 to Trial 6 in the Untrustworthy Condition
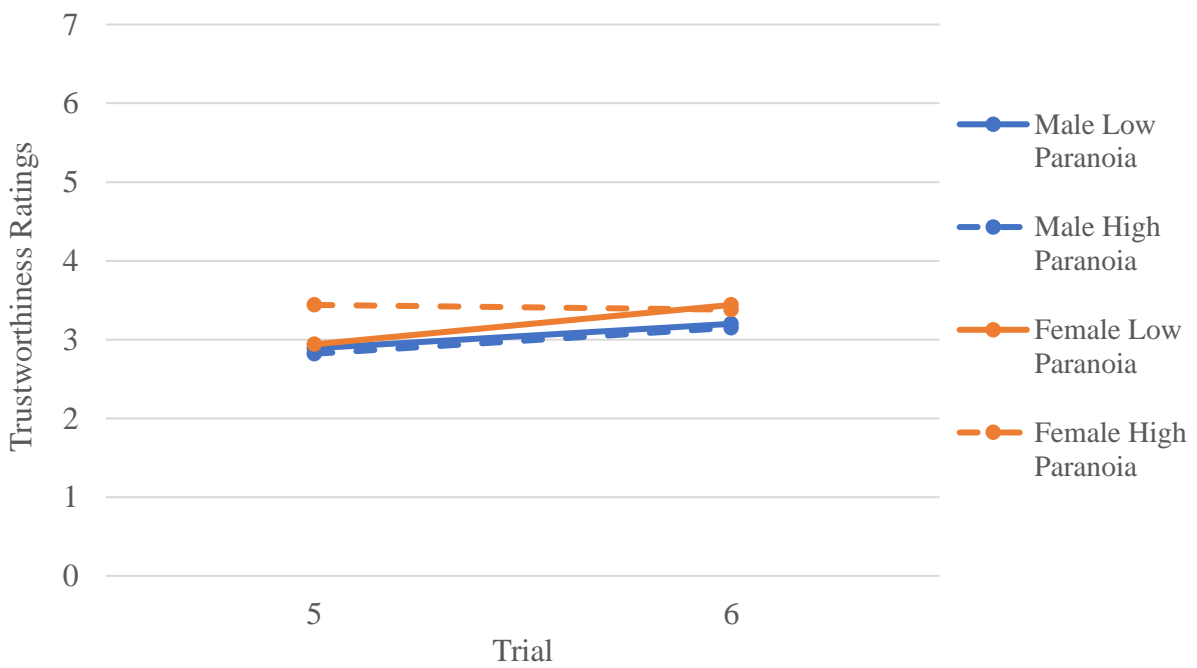
Hypothesis 5

Secondary analysis, using Pearson's correlation was also completed to assess the association between the variables for the low and high paranoia group (Table 5 and Table 6, respectively; for the whole sample see Appendix S).

Low Paranoia

A weak negative association was found between paranoia and self-esteem and attachment anxiety. A moderate negative correlation was found between paranoia and age. All p's < .01. Self-esteem was moderately positively correlated with attachment anxiety (p < .01) and weakly positively correlated with age and the female face difference score in the trustworthy condition (both p's < .05) Attachment anxiety weakly negatively correlated with age (p < .01), but weakly positively correlated with attachment avoidance and the female face difference score in the trustworthy condition (both p's < .05).

A weak positive association was found between age and male face baseline score in the trustworthy condition and the female face difference score in the untrustworthy condition (both p's < .05). Baseline male and female faces from the trustworthy and untrustworthy condition yielded a weak positive correlation with each other (all p's < .05). A weak negative correlation was found for female face trustworthy difference score and male face untrustworthy difference score (p < .05). A weak positive association was found between the male face difference score and the female face difference score in the trustworthy condition (p < .05). All other associations between the remaining variables were non-significant (all p's > .05).

High Paranoia

A weak negative association was found between paranoia and the male face baseline score in the untrustworthy condition ($p < .05$). Self-esteem was moderately positively correlated with attachment anxiety ($p < .01$) as well as attachment avoidance and age (both *p's* $< .05$). Attachment anxiety weakly positively correlated with the male face baseline score in the trustworthy condition ($p < .01$). A weak negative association was found between JTC and the male face baseline score in the trustworthy condition ($p < .05$).

Baseline male and female faces from the trustworthy and untrustworthy condition yielded a weak positive correlation between the male faces in both conditions, the female faces in both conditions, and the female face in the trustworthy condition and the male face in the untrustworthy condition (all *p's* $< .05$). A weak positive association was also found between the male face baseline score and the male face difference score in the trustworthy condition ($p < .05$). All other associations between the remaining variables were non-significant (all *p's* $> .05$).

Table 5

Correlations Between the Variables for the Low Paranoia Group

								Baseline				Difference Scores			
								T		UT		T		UT	
Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. GPTS Total	1	**-.39****	**-.26****	-.05	.06	.08	**-.41****	.02	.02	.06	.06	-.03	-.15	-.10	-.09
2. SE Total		1	**.44****	.14	.03	-.18	**.23***	.19	.08	-.10	-.05	-.04	**.25***	.17	-.15
3. RQ Attachment Anxiety			1	**.23***	-.11	-.13	**-.30****	.16	.11	-.05	-.08	-.10	**.24***	.12	.10
4. RQ Attachment Avoidance				1	.02	.06	-.17	.04	-.02	-.00	.11	.14	-.01	.03	-.08
5. CRT Total					1	.14	-.09	.04	-.09	.00	-.09	-.01	.09	-.08	.01
6. JTC Draws to Decision ($n = 97$)						1	-.08	.11	.05	.12	.17	-.10	-.06	-.10	.08
7. Age							1	**.21***	.05	.08	.06	-.04	.09	.16	**.21***
8. Male								1	**.37****	**.22***	**.35****	.05	.09	.03	.03
9. Female									1	**.27****	**.33****	-.19	.00	**-.21***	-.08
10. Male										1	**.28****	-.12	-.12	.02	.09
11. Female											1	-.06	-.07	.08	-.07
12. Male												1	**.25***	.01	-.13
13. Female													1	-.06	-.09
14. Male														1	.15
15. Female															1

Row labels grouped at left: Baseline (T: 8 Male, 9 Female; UT: 10 Male, 11 Female), Difference Scores (T: 12 Male, 13 Female; UT: 14 Male, 15 Female)

** $p < .01$, * $p < .05$

Note. T = Trustworthy Condition, UT = Untrustworthy Condition, GPTS = Green's Paranoia Thought Scale, SE = Self-Esteem, RQ = Relationship Questionnaire, CRT = Cognitive Reflection Test, JTC = Jumping to Conclusion

Table 6

Correlations Between the Variables for the High Paranoia Group

| | | | | | | | | | Baseline | | | | Difference Scores | | | |
| | | | | | | | | | T | | UT | | T | | UT | |
Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. GPTS Total	1	.02	-.13	.07	-.12	.14	-.16	-.17	.05	**-.25***	.10	.19	-.06	.06	.06
2. SE Total		1	**.52***	**.28***	.03	-.19	**.29***	.21	-.07	.10	-.09	.03	.05	-.06	-.12
3. RQ Attachment Anxiety			1	.10	.02	-.18	.21	**.36****	-.14	.05	-.08	-.07	-.06	-.21	-.06
4. RQ Attachment Avoidance				1	-.04	-.07	-.01	.04	-.09	.08	-.06	.00	.01	-.02	.05
5. CRT Total					1	.13	-.09	.07	-.07	.03	.03	.06	-.13	-.16	.08
6. JTC Draws to Decision (*n* = 68)						1	-.10	**-.28***	.15	-.17	-.07	-.20	.02	.21	.13
7. Age							1	.20	.07	.01	-.02	.16	.05	.01	-.04
8. Male								1	.04	**.26***	.13	**.27***	-.04	-.04	.04
9. Female									1	**.27***	**.27***	-.01	.13	.13	.10
10. Male										1	.21	.02	-.02	-.04	-.09
11. Female											1	-.05	-.10	-.05	.06
12. Male												1	-.08	.21	-.07
13. Female													1	-.02	-.04
14. Male														1	.05
15. Female															1

** *p* < .01, * *p* < .05

Note. T = Trustworthy Condition, UT = Untrustworthy Condition, GPTS = Green's Paranoia Thought Scale, SE = Self-Esteem, RQ = Relationship Questionnaire, CRT = Cognitive Reflection Test, JTC = Jumping to Conclusion

Discussion

Summary of Results

The current study aimed to identify whether there is a difference in the rate a person experiencing high paranoia adjust their estimations of trustworthiness in the light of new information compared to a person experiencing low paranoia. Trustworthy judgement scores at baseline did not differ by group, suggesting both the low and high paranoia groups provided similar trustworthy judgements for the gendered faces in both conditions.

The test of the first hypothesis, that people experiencing high paranoia will be quicker to learn that neutral faces are untrustworthy due to their previously demonstrated bias was non-significant, without even an indication of a trend in the expected direction. This would suggest there is no difference in the change in trustworthiness judgements between groups, even when presented with scenario outcomes that were initially untrustworthy.

Similarly, the test of the second hypothesis, that people experiencing high paranoia will be slower to learn that neutral faces are trustworthy due to their previously demonstrated mistrust bias failed to reach significance, again suggesting there is no difference in the change in trustworthiness judgements between groups, even when presented with scenario outcomes that were initially trustworthy. Overall, the different conditions yielded similar ratings of trustworthiness judgements in both groups, and our hypotheses were not supported.

Kirk et al. (2013) found people experiencing paranoia rated people they had met for the first time as less trustworthy. However, this was not observed in the current study. One difference between the two studies is that Kirk et al. (2013) used faces that had been manipulated to vary in their level of apparent trustworthiness whereas the present study used only neutral faces.

The lack of difference in trustworthy judgements between the two groups in the current study may be explained by this difference. For example, the scenarios within the untrustworthy conditions may not have elicited a greater perceived threat for the high paranoia group and therefore, no difference in ratings between the two groups. However, it is notable that both groups adjusted their judgements of trust in response to updated information about the behaviour of the characters in the scenario narratives, both when the characters consistently behaved as trustworthy or untrustworthy. Hence, this is evidence that the current paradigm was a valid and sensitive measure of trust judgements.

The third and fourth hypotheses were that when encountering a trust violation (somebody who has behaved in a trustworthy fashion suddenly behaves in an untrustworthy fashion or vice-versa), people experiencing high paranoia will show greater reductions in trustworthiness judgements whereas, in the low paranoia group, these judgements will be less affected. However, the interactions between group and time in the relevant analyses were non-significant, therefore not supporting these hypotheses. Although both groups expectedly changed their estimates of trustworthiness (there was a significant effect of time) they did not differ in the way they rated trustworthiness.

Although the gender of the faces did not pertain to the main hypothesis. The current study found a significant main effect of the gender of faces in both conditions and both groups as expected. However, unexpectedly, a significant interaction was found between the gender of face and time in both conditions. This would suggest male and female gendered faces in both conditions were rated differently across time. Female faces were consistently rated more trustworthy compared to male faces and this was evident even when female faces were presented as untrustworthy. These findings are consistent with previous research which noted female faces are more trustworthy compared to male faces. Specifically, gendered faces that were stereotypically female are rated more positively compared to masculine female

faces and male counterparts (Sutherland et al., 2015). However, as the gender analyses in the current study were secondary and not included in the main power analysis the current findings should be interpreted with caution.

The secondary hypothesis that people in the high paranoia group will present with attachment insecurity and low levels of self-esteem and these will predict changes in trustworthiness judgements observed in the two groups was partially supported by the current findings. In support of the hypothesis, the low paranoia group showed higher levels of self-esteem and lower levels of attachment anxiety. However, in the high paranoia group, paranoia did not significantly correlate with lower self-esteem or attachment anxiety or avoidance.

The current findings are inconsistent with the literature which has found a strong association between paranoia and attachment and self-esteem (Lavin et al., 2020; Pickering et al., 2008; Wicham et al., 2015). A recent meta-analysis reviewing 26 studies concluded that both attachment anxiety and attachment avoidance are significantly associated with paranoia and the strength of the association is similar for both clinical and non-clinical populations (Murphey et al., 2020). Similarly, self-esteem is considered to predict the experience of paranoia, with lower levels of self-esteem associated with greater paranoia in both clinical (Freeman et al., 1998) and non-clinical samples (Combs & Penn, 2004; Ellett et al., 2003; Martin & Penn, 2001).

One possible hypothesis about the lack of association between paranoia and both low self-esteem and attachment anxiety in the high paranoia group, in contrast to significant associations in the low paranoia group, is that there was insufficient variance in the paranoia scores of the high group to detect a difference. However, inspection of table 1 indicates that the variance of paranoia was higher in the high group compared to the low group, so this

cannot be the explanation. Notably, between-group comparisons showed lower self-esteem and higher attachment anxiety in the high group, nonetheless. Hence, the probable explanation is either that the relationships between paranoia and both self-esteem and attainment anxiety are non-linear or that there were floor/ceiling effects for the relevant measures (the scales were insufficiently sensitive at the extremes).

Interestingly, the current study found a significant correlation between age and self-esteem, with higher self-esteem in older people for both groups. This would suggest that older people have greater levels of self-esteem compared to younger people. Several reviews, see Orth and Robins (2014), Robins and Trzesniewski (2005), and Robins et al. (2002) have assessed self-esteem across different ages (9-99 years) for a broad range of populations. Repeatedly, self-esteem was shown to be high in childhood, lower in adolescence, and becomes higher as an individual progresses through adulthood. Differences in maturity, attitudes about oneself, developmental changes, and/or societal expectations could explain the difference in self-esteem for older and younger people (Ogihara & Kusumi, 2020).

Within the low paranoia group, age was significantly associated with higher paranoia, where younger people compared to older people experienced greater paranoia severity. Increased paranoia in non-clinical young people may relate to their social experiences, which are likely to occur online (Ogihara & Kusumi, 2020). Previous research suggests that elevated use of online platforms and heightened emotional reactivity to social media content may increase feelings of vulnerability or threat in young people, and therefore may influence the likelihood of paranoia (Guglielmucci et al., 2017; Xiuqin et al., 2010). Further research examining the relationship between these variables may be clinically beneficial in the management of paranoia in young people.

Strengths and Limitations

The current study is the first to implement a novel task to assess dynamic updating of levels of trust for people experiencing low and high levels of paranoia, therefore, the findings can be considered to have greater ecological validity.

Additionally, despite not finding a significant difference in the rate a person with low or high paranoia learns trustworthiness, the study can be commended for developing a successful paradigm that measures the level of trustworthiness across time.

Moreover, efforts were made to increase the diversity of the sample by recruiting participants through multiple platforms (social media sites and Prolific). However, most of the sample were white British and female which does not represent a diverse sample.

Another limitation was the missed JTC task from the main study for the high paranoia group. This led to a 13% dropout rate; however, it was deemed to be low. Finally, an oversight for the current study was the lack of attention bias tasks during the computerised study. In hindsight, including these tasks may have yielded greater validity in the results.

Clinical and Theoretical Implications and Future Research

The current findings contribute to the literature and understanding of paranoia and interpersonal trust as the study adopted an ecologically valid task. Consequently, the conclusions drawn from findings could be considered to have greater validity compared to previous studies. Though, the results indicated that people experiencing low and high paranoia do not differ in the way they update their trust judgments.

While the study findings were largely negative, the null hypothesis could not be rejected, the role of trust in clinical practice with patients experiencing paranoia merits close study. The issue of trust is likely to be important in working with people experiencing

paranoia. As a result, practitioners are likely to improve their ability to form therapeutic alliances if they pay attention to the way trust evolves within therapeutic sessions.

Future research would benefit from assessing the differences in trustworthiness between clinical and subclinical samples using the study paradigm. This will help conclude whether a difference between the low and high paranoia group was not detected due to a small and/or non-clinical sample.

Conclusions

The current study is the first to use a novel and dynamic approach to assess the difference in the rate a person experiencing low or high paranoia adjusts their estimations of trustworthiness in the light of new information. The findings indicated a non-significant difference in the rate a person changes trustworthiness judgements between the groups. However, trust judgements changed in response to additional information in an expected way, suggesting that the paradigm is sensitive to changing estimates of trustworthiness. In both conditions, there was a significant interaction between gender and time, with greater trustworthiness rating for females across time, and between age and group, with greater paranoia severity demonstrated in younger people. However, inconsistent with previous findings, attachment anxiety and avoidance as well as self-esteem did not significantly correlate with the high paranoia. Future research may benefit from using the current paradigm, with clinical and non-clinical samples to further the understanding of paranoia and trustworthiness.

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21-62. https://doi.org/10.1177/1534582302001001003

Baas, D., Van't Wout, M., Aleman, A., & Kahn, R. S. (2008). Social judgement in clinically stable patients with schizophrenia and healthy relatives: Behavioural evidence of social brain dysfunction. *Psychological Medicine*, *38*(5), 747-754. https://doi.org/10.1017/S0033291707001729

Bartholomew, K. (1990). Avoidance of intimacy: An attachment perspective. *Journal of Social and Personal Relationships*, *7*(2), 147-178. https://doi.org/10.1177/0265407590072001

Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology, 61*(2), 226–244. https://doi.org/10.1037/0022-3514.61.2.226

Bebbington, P. E., McBride, O., Steel, C., Kuipers, E., Radovanoviĉ, M., Brugha, T., Jenkins, R., Meltzer, H. I., & Freeman, D. (2013). The structure of paranoia in the general population. *The British Journal of Psychiatry*, *202*(6), 419-427.

Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory Delusions: A Review and Theoretical Integration. *Clinical Psychology Review*, *21*(8), 1143–1192. https://doi.org/10.1016/S0272-7358(01)00106-4

Bentall, R. P., de Sousa, P., Varese, F., Wickham, S., Sitko, K., Haarmans, M., & Read, J. (2014). From adversity to psychosis: pathways and mechanisms from specific adversities to specific symptoms. *Social Psychiatry and Psychiatric Epidemiology*, *49*(7), 1011-1022. doi: 10.1007/s00127-014-0914-0

Bentall, R. P., & Fernyhough, C. (2008). Social predictors of psychotic experiences: specificity and psychological mechanisms. *Schizophrenia Bulletin*, *34*(6), 1012-1020. https://doi.org/10.1093/schbul/sbn103

Bentall, R. P., Rowse, G., Shryane, N., Kinderman, P., Howard, R., Blackwood, N., ... & Corcoran, R. (2009). The cognitive and affective structure of paranoid delusions: a transdiagnostic investigation of patients with schizophrenia spectrum disorders and depression. *Archives of General Psychiatry*, *66*(3), 236-247. doi:10.1001/archgenpsychiatry.2009.1

Bronstein, M. V., Everaert, J., Castro, A., Joormann, J., & Cannon, T. D. (2019). Pathways to paranoia: Analytic thinking and belief flexibility. *Behaviour Research and Therapy*, *113*, 18-24. https://doi.org/10.1016/j.brat.2018.12.006

Bowlby J. (1969). *Attachment. Attachment and loss: Vol. 1. Loss*. New York: Basic Books

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, 2nd edn. Á/L.

Colbert, S. M., & Peters, E. R. (2002). Need for closure and jumping-to-conclusions in delusion-prone individuals. *The Journal of Nervous and Mental Disease*, *190*(1), 27-31. doi: 10.1097/00005053-200201000-00007

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of applied psychology*, *92*(4), 909. https://doi.org/10.1037/0021-9010.92.4.909

Combs, D. R., & Penn, D. L. (2004). The role of subclinical paranoia on social perception and behavior. *Schizophrenia Research*, *69*(1), 93-104. https://doi.org/10.1016/S0920-9964(03)00051-3

Dudley, R., Daley, K., Nicholson, M., Shaftoe, D., Spencer, H., Cavanagh, K., & Freeston, M. (2013). 'Jumping to conclusions' in first-episode psychosis: A longitudinal study. *British Journal of Clinical Psychology*, *52*(4), 380-393. https://doi.org/10.1111/bjc.12023

Elahi, A., Algorta, G. P., Varese, F., McIntyre, J. C., & Bentall, R. P. (2017). Do paranoid delusions exist on a continuum with subclinical paranoia? A multi-method taxometric study. *Schizophrenia Research*, *190*, 77-81. doi: 10.1016/j.schres.2017.03.022

Ellett, L. Y. N., Lopes, B., & Chadwick, P. (2003). Paranoia in a nonclinical population of college students. *The Journal of Nervous and Mental Disease*, *191*(7), 425-430. doi: 10.1097/01.NMD.0000081646.33030.EF.

Evans, J. D. (1996). *Straightforward Statistics for the Behavioral Sciences*. Thomson Brooks/Cole Publishing Co.

Fine, C., Gardner, M., Craigie, J., & Gold, I. (2007). Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cognitive Neuropsychiatry*, *12*(1), 46-77. https://doi.org/10.1080/13546800600750597

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25-42. doi: 10.1257/089533005775196732

Freeman, D., Evans, N., & Lister, R. (2012). Gut feelings, deliberative thought, and paranoid ideation: A study of experiential and rational reasoning. *Psychiatry Research*, *197*(1-2), 119-122. https://doi.org/10.1016/j.psychres.2011.12.031

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., … Dunn, G. (2005). Psychological investigation of the structure of paranoia. *The British Journal of Psychiatry*, *186*(5). https://doi.org/10.1192/bjp.186.5.427

Freeman, D., Garety, P., Fowler, D., Kuipers, E., Dunn, G., Bebbington, P., & Hadley, C. (1998). The London-East Anglia randomized controlled trial of cognitive-behaviour therapy for psychosis IV: Self-esteem and persecutory delusions. *British Journal of Clinical Psychology*, *37*(4), 415-430. https://doi.org/10.1111/j.2044-8260.1998.tb01399.x

Freeman, D., Lister, R., & Evans, N. (2014b). The use of intuitive and analytic reasoning styles by patients with persecutory delusions. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(4), 454-458. https://doi.org/10.1016/j.jbtep.2014.06.005

Freeman, D., Loe, B. S., Kingdon, D., Startup, H., Molodynski, A., Rosebrock, L., ... & Bird, J. C. (2019). The revised Green et al., Paranoid Thoughts Scale (R-GPTS): psychometric properties, severity ranges, and clinical cut-offs. *Psychological Medicine*, 1-10. https://doi.org/10.1017/S0033291719003155

Freeman, D., Pugh, K., & Garety, P. (2008). Jumping to conclusions and paranoid ideation in the general population. *Schizophrenia Research*, *102*(1-3), 254-260. doi:10.1016/j.schres.2008.03.020

Freeman, D., Pugh, K., Vorontsova, N., Antley, A., & Slater, M. (2010). Testing the continuum of delusional beliefs: an experimental study using virtual reality. *Journal of Abnormal Psychology*, *119*(1), 83-92. http://dx.doi.org/10.1037/a0017514

Freeman, D., Startup, H., Dunn, G., Černis, E., Wingham, G., Pugh, K., ... & Kingdon, D. (2014a). Understanding jumping to conclusions in patients with persecutory delusions: working memory and intolerance of uncertainty. *Psychological Medicine*, *44*(14), 3017-3024. doi:10.1017/S0033291714000592

Garety, P. A., Freeman, D., Jolley, S., Dunn, G., Bebbington, P. E., Fowler, D. G., ... & Dudley, R. (2005). Reasoning, emotions, and delusional conviction in

psychosis. *Journal of Abnormal Psychology*, *114*(3), 373.

https://doi.org/10.1037/0021-843X.114.3.373

Garety, P., Joyce, E., Jolley, S., Emsley, R., Waller, H., Kuipers, E., ... & Freeman, D.

(2013). Neuropsychological functioning and jumping to conclusions in

delusions. *Schizophrenia Research*, *150*(2-3), 570-574.

https://doi.org/10.1016/j.schres.2013.08.035

Green, M. F., Penn, D. L., Bentall, R., Carpenter, W. T., Gaebel, W., Gur, R. C., … Heinssen,

R. (2008a). Social Cognition in Schizophrenia: An NIMH Workshop on Definitions,

Assessment, and Research Opportunities. *Schizophrenia Bulletin*, *34*(6), 1211–1220.

https://doi.org/10.1093/schbul/sbm14

Green, C. E. L., Freeman, D., Kuipers, E., Bebbington, P., Fowler, D., Dunn, G., & Garety, P.

A. (2008b). Measuring ideas of persecution and social reference: The Green et al.

Paranoid Thought Scales (GPTS). *Psychological Medicine*, *38*(1), 101-111.

doi: 10.1017/S0033291707001638

Griffin, D. W., & Bartholomew, K. (1994). The metaphysics of measurement: The case of

adult attachment. In K. Bartholomew & D. Perlman (Eds.), *Advances in personal*

relationships: Vol. 5. Attachment processes in adulthood (pp. 17-52). Jessica

Kingsley.

Guglielmucci, F., Saroldi, M., Zullo, G., Munno, D., & Granieri, A. (2017). Personality

profiles and problematic internet use in a sample of Italian adolescents. *Clinical*

Neuropsychiatry, 14(1), 94-103.

Hajdúk, M., Krajčovičová, D., Zimányiová, M., Kořínková, V., Heretik, A., & Pečeňák, J.

(2019). Trustworthiness judgments are associated with interpersonal functioning in

individuals with schizophrenia. *Journal of Experimental Psychopathology*, *10*(1), 1-9. https://doi.org/10.1177/2043808719827094

Hassan, Z. A., Schattner, P., & Mazza, D. (2006). Doing a pilot study: why is it essential?. *Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia*, *1*(2-3), 70. PMID: 27570591

Humphrey, C., Bucci, S., Varese, F., Degnan, A., & Berry, K. (2021). Paranoia and negative schema about the self and others: A systematic review and meta-analysis. *Clinical Psychology Review*, *90*, 102081. https://doi.org/10.1016/j.cpr.2021.102081

Kaney, S., Wolfenden, M., Dewey, M. E., & Bentall, R. P. (1992). Persecutory delusions and recall of threatening propositions. *British Journal of Clinical Psychology*, *31*(1), 85-87. https://doi.org/10.1111/j.2044-8260.1992.tb00971.x

Kim, S. H., Baek, M., & Park, S. (2021). Association of Parent–child Experiences with Insecure Attachment in Adulthood: A Systematic Review and Meta-analysis. *Journal of Family Theory & Review*, *13*(1), 58-76. https://doi.org/10.1111/jftr.12402

Kirk, H., Gilmour, A., Dudley, R., & Riby, D. (2013). Paranoid Ideation and Assessments of Trust. *Journal of Experimental Psychopathology*, *4*(4), 360–367. https://doi.org/10.5127/jep.027812

Lavin, R., Bucci, S., Varese, F., & Berry, K. (2020). The relationship between insecure attachment and paranoia in psychosis: A systematic literature review. *British Journal of Clinical Psychology*, *59*(1), 39-65. doi: 10.1111/bjc.12231

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

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response

    categories on the reliability and validity of rating scales. *Methodology*, *4*(2), 73-79.

    https://doi.org/10.1027/1614-2241.4.2.73

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces

    (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section,*

    Karolinska Institutet, *91*(630), 2-2.

MacBeth, A., Gumley, A., Schwannauer, M., & Fisher, R. (2011). Attachment states of mind,

    mentalization, and their correlates in a first-episode psychosis sample. *Psychology*

    and Psychotherapy: Theory, Research and Practice, *84*(1), 42-57.

    https://doi.org/10.1348/147608310X530246

Martin, J. A., & Penn, D. L. (2001). Social cognition and subclinical paranoid ideation.

    British Journal of Clinical Psychology, *40*(3), 261-265.

    https://doi.org/10.1348/014466501163670

Martinez, A. P., Agostini, M., Al-Suhibani, A., & Bentall, R. P. (2021). Mistrust and negative

    self-esteem: Two paths from attachment styles to paranoia. *Psychology and*

    Psychotherapy: Theory, Research and Practice, *94*(3), 391-406.

    https://doi.org/10.1111/papt.12314

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of

    organizational trust. *Academy of Management Review*, *20*(3), 709-734.

    https://doi.org/10.2307/258792

McKay, R., Langdon, R., & Coltheart, M. (2006). Need for closure, jumping to conclusions,

    and decisiveness in delusion-prone individuals. *The Journal of Nervous and Mental*

    Disease, *194*(6), 422-426. doi: 10.1097/01.nmd.0000221353.44132.25

Menon, M., Mizrahi, R., & Kapur, S. (2008). 'Jumping to conclusions' and delusions in psychosis: relationship and response to treatment. *Schizophrenia Research*, *98*(1-3), 225-231. https://doi.org/10.1016/j.schres.2007.08.021

Moore, C. G., Carter, R. E., Nietert, P. J., & Stewart, P. W. (2011). Recommendations for planning pilot studies in clinical and translational research. *Clinical and Translational Science*, *4*(5), 332-337. doi: 10.1111/j.1752-8062.2011.00347.x.

Murphy, R., Goodall, K., & Woodrow, A. (2020). The relationship between attachment insecurity and experiences on the paranoia continuum: A meta-analysis. *British Journal of Clinical Psychology*, *59*(3), 290-318. https://doi.org/10.1111/bjc.12247.

Ogihara, Y., & Kusumi, T. (2020). The developmental trajectory of self-esteem across the life span in Japan: age differences in scores on the Rosenberg Self-esteem scale from adolescence to old age. *Frontiers in Public Health*, *8,* 132. https://doi.org/10.3389/fpubh.2020.00132

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science*, *23*(5), 381-387. https://doi.org/10.1177/0963721414547414

Oxford English Dictionary (1989). Oxford English Dictionary. *Simpson, JA & Weiner, ESC*.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, *72*(3), 346-354. https://doi.org/10.1037/h0023653.

Phillips, M. L., Senior, C., & David, A. S. (2000). Perception of threat in schizophrenics with persecutory delusions: an investigation using visual scan paths. *Psychological Medicine*, *30*(1), 157-167. https://doi.org/10.1017/S0033291799001397

Pickering, L., Simpson, J., & Bentall, R. P. (2008). Insecure attachment predicts proneness to paranoia but not hallucinations. *Personality and Individual Differences*, *44*(5), 1212-1224. https://doi.org/10.1016/j.paid.2007.11.016

Reed, E. J., Uddenberg, S., Suthaharan, P., Mathys, C. D., Taylor, J. R., Groman, S. M., & Corlett, P. R. (2020). Paranoia as a deficit in non-social belief updating. *eLife*, *9*, e56345. https://doi.org/10.7554/eLife.56345

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49*(1), 95. https://doi.org/10.1037/0022-3514.49.1.95

Ringer, J. M., Buchanan, E. E., Olesek, K., & Lysaker, P. H. (2014). Anxious and avoidant attachment styles and indicators of recovery in schizophrenia: Associations with self-esteem and hope. *Psychology and Psychotherapy: Theory, Research and Practice*, *87*(2), 209-221. doi: 10.1111/papt.12012

Robins, R. W., & Trzesniewski, K. H. (2005). Self-esteem development across the lifespan. *Current Directions in Psychological Science*, *14*(3), 158-162. https://doi.org/10.1111/j.0963-7214.2005.00353.x

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17*(3), 423-434. https://doi.org/10.1037/0882-7974.17.3.423

Ross, R. M., McKay, R., Coltheart, M., & Langdon, R. (2015). Jumping to conclusions about the beads task? A meta-analysis of delusional ideation and data-gathering. *Schizophrenia Bulletin*, *41*(5), 1183-1191. https://doi.org/10.1093/schbul/sbu187

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393-404. doi: 10.5465/AMR.1998.926617

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.

Sitko, K., Bentall, R. P., Shevlin, M., & Sellwood, W. (2014). Associations between specific psychotic symptoms and specific childhood adversities are mediated by attachment styles: an analysis of the National Comorbidity Survey. *Psychiatry Research*, *217*(3), 202-209. doi: 10.1016/j.psychres.2014.03.019

Speechley, W. J., Murray, C. B., McKay, R. M., Munz, M. T., & Ngan, E. T. C. (2010). A failure of conflict to modulate dual-stream processing may underlie the formation and maintenance of delusions. *European Psychiatry*, *25*(2), 80-86. doi:10.1016/j.eurpsy.2009.05.012

Speechley, W. J., Woodward, T. S., & Ngan, E. T. (2013). Failure of conflict to modulate central executive network activity associated with delusions in schizophrenia. *Frontiers in Psychiatry*, *4*, 113. https://doi.org/10.3389/fpsyt.2013.00113

Startup, H., Freeman, D., & Garety, P. A. (2008). Jumping to conclusions and persecutory delusions. *European Psychiatry*, *23*(6), 457-459. doi: 10.1016/j.eurpsy.2008.04.005

Sutherland, C. A., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology*, *106*(2), 186-208. doi: 10.1111/bjop.12085.

Thewissen, V., Bentall, R. P., Oorschot, M., A Campo, J., van Lierop, T., van Os, J., & Myin-Germeys, I. (2011). Emotions, self-esteem, and paranoid episodes: an experience sampling study. *The British Journal of Clinical Psychology*, *50*(2), 178–195. https://doi.org/10.1348/014466510X508677

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99. doi: 10.1037/t49856-000

Trotta, A., Kang, J., Stahl, D., & Yiend, J. (2020). Interpretation bias in paranoia: A systematic review and meta-analysis. *Clinical Psychological Science*, doi: 2167702620951552.

Van Dael, F., Versmissen, D., Janssen, I., Myin-Germeys, I., Van Os, J., & Krabbendam, L. (2006). Data gathering: biased in psychosis?. *Schizophrenia Bulletin*, *32*(2), 341-351. https://doi.org/10.1093/schbul/sbj021

Ward, T., & Garety, P. A. (2019). Fast and slow thinking in distressing delusions: A review of the literature and implications for targeted therapy. *Schizophrenia Research*, *203*, 80-87. https://doi.org/10.1016/j.schres.2017.08.045

Ward, T., Peters, E., Jackson, M., Day, F., & Garety, P. A. (2018). Data-gathering, belief flexibility, and reasoning across the psychosis continuum. *Schizophrenia Bulletin*, *44*(1), 126-136. https://doi.org/10.1093/schbul/sbx029

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, *27*(3), 236-247. https://doi.org/10.1016/j.ijresmar.2010.02.004

Wickham, S., Sitko, K., & Bentall, R. P. (2015). Insecure attachment is associated with paranoia but not hallucinations in psychotic patients: the mediating role of negative self-esteem. *Psychological Medicine*, *45*(7), 1495. doi:10.1017/S0033291714002633

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592-598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wincenciak, J., Dzhelyova, M., Perrett, D. I., & Barraclough, N. E. (2013). Adaptation to facial trustworthiness is different in female and male observers. *Vision research*, *87*, 30-34. doi: 10.1016/j.visres.2013.05.007

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*(3), 277-283. doi: 10.1038/nn816

Xiuqin, H., Huimin, Z., Mengchen, L., Jinan, W., Ying, Z., & Ran, T. (2010). Mental health, personality, and parental rearing styles of adolescents with internet addiction disorder. *Cyberpsychology, Behavior and Social Networking*, *13*(4), 401–406. https://doi.org/10.1089/cyber.2009.0222

Zak, P. J., & Knack, S. (2001). Trust and growth. *The Economic Journal*, *111*(470), 295-321. https://doi.org/10.1111/1468-0297.00609

Appendices

Appendix A. Green's Paranoia Thought Scale - B

Redacted due to copyright

Appendix B. Participant Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
• Non-clinical adult sample	• Participants who do not consent to take part in the study
• Participants aged 18 and over	
• Males and Females	• Missing demographic data
• Participants who have completed the Green's Paranoid Thought Scale	• Missing data from the computerised task
• Read and Understand English	• Participants with a score of 34 and below on the Green's Paranoid Thought Scale when recruited via Prolific

Appendix C. Self-Esteem Rating Scale - Short Form

Redacted due to copyright

Appendix D. Relationship Questionnaire

Redacted due to copyright

Appendix E. Cognitive Reflection Task

Redacted due to copyright

Appendix F. Jumping to Conclusions Task

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Redacted due to copyright

Appendix G. Faces used in the Paradigm

Male and Female Trustworthy Condition

Male and Female Untrustworthy Condition

Appendix H. Scenarios in the Paradigm

Trust Level	Condition 1 (Male 1) Trustworthy – Violates	Condition 2 (Male 2) Untrustworthy – Violates	Condition 3 (Female 1) Trustworthy – Violates	Condition 4 (Female 2) Untrustworthy – Violates
	Transport	*Colleague*	*Neighbour*	*Stranger*
LOW				
Scenario 1	You are standing next to this person who is also waiting for the train. You realise you have left your bank card in the ticket office.	You have had to leave work early for an appointment. You ask this colleague to give a file to your boss.	You notice this person is your new neighbour. You are going to be working late and are expecting a parcel.	You are driving in a new city and are unsure about how to get to the town centre. You see this person nearby.
Question	How trustworthy would you say this person is to watch your luggage whilst you collect your bank card?	How trustworthy would you say this person is to deliver the file to your boss?	How trustworthy would you say this person is to collect your parcel?	How trustworthy would you say this person is to give you reliable directions to the town centre?
Outcome	You return from the ticket office and see that the person has watched your luggage, and nothing is missing.	You got to work the next day and see the file is still on your desk. Your colleague did not hand the file to your boss.	You receive an email confirming your parcel has been successfully delivered. When you return home your neighbour brought round the parcel.	You followed the directions you were given but you realise partway through the journey that they are wrong. You have ended up in a different part of the town.
Scenario 2	You get on the train and find yourself sitting near this person. You want to get some food, but your phone needs to charge as you have an email ticket.	You are called away to an emergency. You ask this colleague to inform your team that you have had to leave.	You have forgotten to take the bins out for collection before leaving for work. You phone your neighbour to ask if they could take your bins out for you.	You are struggling to reverse park your car and notice this person is guiding you into the bay.

Trust Level	Condition 1 (Male 1) Trustworthy – Violates	Condition 2 (Male 2) Untrustworthy – Violates	Condition 3 (Female 1) Trustworthy – Violates	Condition 4 (Female 2) Untrustworthy – Violates
	Transport	*Colleague*	*Neighbour*	*Stranger*
Question	How trustworthy would you say this person is to watch your phone whilst it is on charge?	How trustworthy would you say this person is to notify your team that you had to leave?	How trustworthy would you say this person is to put your bins out for collection?	How trustworthy would you say this person is to help you reverse park your car into a bay?
Outcome	You return to your seat with your food and see that the person has watched your phone whilst it was charging.	You return to work and your team ask you where you had been all morning. You realise your colleague did not inform your team that you had to leave for an emergency.	You return home and notice your bins have moved and that your neighbour has taken the bins out for you.	You decided to follow the guidance but when you step out of the car you realise that you have parked very badly, and your car is touching another car.
Scenario 3	The train has come to a standstill, and you have arrived at the airport. You want to send a picture with the sign to your friend and you see this person standing near you.	You have been called away to another job. You ask this colleague to switch your laptop off as confidential information is visible.	Your car is being repaired at a garage. You ask this neighbour for a lift to work tomorrow.	You want to see the main attraction in the town centre, but you are unsure of the opening times. You see this person nearby.
Question	How trustworthy would you say this person is to take a picture with your phone?	How trustworthy would you say your colleague is to switch off your laptop?	How trustworthy would you say your neighbour is to give you a lift to work tomorrow?	How trustworthy would you say this person is to give you the correct information about the opening times?

Trust Level	Condition 1 (YY) Trustworthy – Constant	Condition 2 (NN) Untrustworthy – Constant	Condition 3 (YN) Trustworthy – Violates	Condition 4 (NY) Untrustworthy – Violates
	Transport	*Colleague*	*Neighbour*	*Stranger*
Outcome	This person agrees to take a picture of you and helped you to pose for the picture. They then returned your phone.	After completing the job you return to your desk. You see that your colleague has not switched your laptop off, and the confidential information remains visible.	Your neighbour came over to your house in the morning and waited to give you a lift to work.	The person tells you that the attraction is open for another two hours but when you arrive the attraction is closed. The times you were given were incorrect.
MEDIUM				
Scenario 4	You have cleared airport security and are in the departures lounge waiting for your flight. You try to connect to the airport Wi-Fi, but you are having trouble logging in with your details. You see this person a few seats away.	You are going on annual leave and ask this colleague to add your share of the money to the monthly milk kitty.	You are having to work late. You have asked this neighbour to put food out for your cat at 5pm.	You are meeting a friend at a café, but you are unsure which tram line you need to take. You see this person waiting for a tram too.
Question	How trustworthy would you say this person is with your details to help you access the airport Wi-Fi?	How trustworthy would you say this person is to add your share of the money to the kitty?	How trustworthy would you say this person is to feed your cat?	How trustworthy would you say this person is to give you the correct tram line?
Outcome	They help you to access the Wi-Fi at the airport and then show you how to delete your details from the website.	You return from your annual leave and see that your colleague did not add your share of the money to the kitty.	Your neighbour messages you at 5.05pm to let you know that they have fed your cat.	The instructions you were given are for the wrong tram line and you are delayed in meeting your friend.

Trust Level	Condition 1 (YN - Male) Trustworthy – Violation	Condition 2 (YN - Female) Untrustworthy – Violation	Condition 3 (YN - Male) Trustworthy – Violates	Condition 4 (NY- Female) Untrustworthy – Violates
	Transport	*Colleague*	*Neighbour*	*Stranger*
Scenario 5 (Outcome violation trials)	Whilst waiting for your flight you decide to buy some souvenirs. There is a restriction of two per customer, but you need to buy three. You ask this person to buy you one of the souvenirs with your money.	You are running late and need to deliver an important presentation. You ask this colleague to start the presentation without you.	You are going away on holiday and ask this neighbour to water your outdoor plants whilst you are away.	You want to go into a shop to buy some milk, but you cannot take your dog in. You see this person exiting the shop.
Question	How trustworthy would you say this person is with your money to buy the souvenir on your behalf?	How trustworthy would you say this person is to start delivering the presentation for you?	How trustworthy would you say this person is to water your outdoor plants?	How trustworthy would you say this person is to watch your dog?
Outcome (Outcome violation)	You gave this person £20 to purchase the souvenir that costs £15. They brought you the souvenir but did not give you back your change.	You arrive at the meeting and see that your colleague has started the presentation exactly as you asked.	You return from holiday after a week. You see that your plants have not survived as your neighbour has forgotten to water your plants.	The person looked after your dog and treated it nicely whilst you went into the shop and brought the milk.

Trust Level	Condition 1 (Male 1) Trustworthy – Violates	Condition 2 (Male 2) Untrustworthy – Violates	Condition 3 (Female 1) Trustworthy – Violates	Condition 4 (Female 2) Untrustworthy – Violates
	Transport	*Colleague*	*Neighbour*	*Stranger*
HIGH				
Scenario 6	It is the last call for your flight. You are waiting in line next to this person to board the plane. You realise you have left your jacket in the waiting area.	You are working from home and need to change your computer login details, which can only be done from your work computer. You know this colleague is in the office today.	You are getting your boiler fixed but you will not be at home during the specified time slot. You know this neighbour has a day off today.	You need to pay for the carpark, but you are running late for an important event. You see this person in line for the parking pay machine.
Question	How trustworthy would you say this person is to wait for you with your passport and prevent the gates from closing?	How trustworthy would you say this person is with your password to change your login details?	How trustworthy would you say this person is with a spare set of your house keys to let the engineer in?	How trustworthy would you say this person is with your cash to pay for your parking?
Outcome	You found your jacket and return to the gate. You see that this person has waited for you to return before boarding the plane. They hand you back your passport.	Your colleague takes your password, but you still cannot access your computer. You realise your colleague has not changed your login details for you.	Your neighbour goes to your home early to ensure they are present when the engineer arrives. They let the engineer in, and your boiler is fixed.	You gave cash to the person to pay for the parking, but they pocketed the money. A week later you received a fine from the parking company.

Appendix I. Example Pilot Feedback

Improvements to the study

- Make the font larger.

- Make the gendered faces larger.

- Potentially present the questionnaire i.e. the self-esteem measure in a table to ease rating as opposed to having to click through for each question.

- Reword scenario to "*The train has come to a standstill, and you have arrived at the airport. You want to send a picture with the sign to your friend and you see this person standing near you*" to ease understanding of the scenario.

Feedback

"Really enjoyable task, it made me think about whether I would trust people to do things if they broke my trust.".

Appendix J. Invite Poster

Appendix K. Project Information Sheet - Low Paranoia

The University Of Sheffield.

Project Information Sheet

1. Research Project Title:

An experimental study investigating suspiciousness and interpersonal trust.

2. Invitation

You are being invited to take part in a research project. Before you decide whether to participate, it is important for you to understand why the research is being conducted and what it will involve. Please take your time in reading the following information and considering whether or not you wish to take part. Please contact the researcher using the contact details below if you would like any further information or clarification. We thank you for your careful consideration.

3. What is the project's purpose?

The aim of the current study is to identify whether there is a difference in levels of suspiciousness and the rate a person learns trustworthiness. This research will be used to write a thesis that fulfils part of the doctoral training.

4. Why have I been chosen?

You have been chosen as you fit the sample to investigate the aims of the study. We will aim to recruit 120 participants to take part in the study.

5. Do I have to take part?

You can decide whether to take part in this study. A copy of the information provided here is yours to keep if you do decide to take part. You can still decide to withdraw* during the process of the study without any negative consequences. You do not have to give a reason for your withdrawal. If you wish to withdraw from the research, please contact the lead researcher Bejal Fatania at bfatania1@sheffield.ac.uk.

*Please note that there is a point at which it will not be possible for your data to be withdrawn from the research. This will be upon the submission of your responses as data is collected anonymously and following submission, it is not identifiable for withdrawal.

6. What will happen to me if I take part? What do I have to do?

If you wish to take part in the study you will be asked to complete three short questionnaires of suspiciousness, your self-esteem, and your relationships with others, as well as two reasoning tasks.

Following this, you will complete an online task where you will be asked to read a scenario in which you will meet four people and try to figure out how trustworthy or untrustworthy they are.

The duration of the study from start to finish should approximately take 30 minutes. We will ask you to provide your email address so that you can be entered into a £25 Amazon voucher prize draw. When the study is closed, we will select one random winner, and notify them by email.

7. What are the possible risks of taking part?

There are no foreseeable disadvantages or risks of taking part in the study.

8.   What are the possible benefits of taking part?

By successfully completing the study, you will be placed in a draw for your chance to win a £25 Amazon voucher. It is also hoped that this research will contribute to the current understanding of trust and suspiciousness. Moreover, it may provide further information about how practice within mental health services can be adapted to better support clients experiencing varying levels of suspiciousness.

9.   Will my taking part in this project be kept confidential?

All the information that we collect about you during the research will be kept strictly confidential and will only be accessible to members of the research team. Data will be anonymised and therefore you will not be identified in any reports or publications. If you agree to us sharing the information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included.

10.   What is the legal basis for processing my personal data?

According to data protection legislation, we are required to inform you that the legal basis we are applying in order to process your personal data is that, 'processing is necessary for the performance of a task carried out in the public interest' (Article 6(1)(e)). Further information can be found in the University's Privacy Notice https://www.sheffield.ac.uk/govern/data-protection/privacy/general.

Only the research team will have access to your personal information that identifies you (your email address) to let you know if you win the £25 Amazon voucher prize draw. All email addresses will be encrypted and not shared. After the study has ended, all email addresses will be securely deleted. Your personal information will not be accessed or used at any other times during the research.

11.   What will happen to the data collected and the results of the research project?

Collected data will only be accessed by the research team after it has been anonymised. The anonymised data will be stored on a secure database until the research project has received the final mark. The data will then be archived in the University of Sheffield Online Research Data (ORDA).

The results of the research are likely to be disseminated and published. The final project will be uploaded to White Rose eTheses online where you can request a copy of the study and view the published results. You will not be identifiable in any of the publications.

Due to the nature of this research, other researchers may find the collected data useful in answering future research questions. We will ask for your explicit consent for your data to be shared in this way.

12.   Who is organising and funding the research?

No organisation and/or company is funding the research.

13.   Who is the Data Controller?

The University of Sheffield will act as the Data Controller for this study. This means that the University is responsible for looking after your information and using it properly.

14.   Who has ethically reviewed the project?

This project has been ethically approved via the University of Sheffield's Ethics Review Procedure, as administered by the Psychology department.

The
University
Of
Sheffield.

15.   What if something goes wrong and I wish to complain about the research?

If you would like to make a complaint about this project, in the first instance you should contact the lead researcher or their supervisor on the contact details below.

If you feel that your complaint has not been handled to your satisfaction you can contact Professor Elizabeth Milne, Head of the Psychology Department at e.milne@sheffield.ac.uk or Dr Jilly Gibson-Miller, chair of the Department Ethics Subcommittee on jilly.gibson@sheffield.ac.uk.

If your complaint relates to how your personal data has been handled, information about how to raise a complaint can be found in the University's Privacy Notice: https://www.sheffield.ac.uk/govern/data-protection/privacy/general.

16.   Who can I contact for further information?

Lead Researcher: Bejal Fatania

Email: bfatania1@sheffield.ac.uk

Research Supervisor: Professor Richard Bentall

Email: r.bentall@sheffield.ac.uk

Alternatively, you can email a.sinha@sheffield.ac.uk or leave a telephone message with Amrit Sinha, Research Support Officer on 0114 222 6650 and he will ask the lead researcher to contact you.

I would like to take this opportunity to thank you for your participation in the study.

Appendix L. Project Information Sheet - High Paranoia

The University Of Sheffield.

Project Information Sheet

1. **Research Project Title:**

An experimental study investigating suspiciousness and interpersonal trust.

2. **Invitation**

You are being invited to take part in a research project. Before you decide whether to participate, it is important for you to understand why the research is being conducted and what it will involve. Please take your time in reading the following information and considering whether or not you wish to take part. Please contact the researcher using the contact details below if you would like any further information or clarification. We thank you for your careful consideration.

3. **What is the project's purpose?**

The aim of the current study is to identify whether there is a difference in levels of suspiciousness and the rate a person learns trustworthiness. This research will be used to write a thesis that fulfils part of the doctoral training.

4. **Why have I been chosen?**

Study participants will be selected following the completion of the initial questionnaire. You have been chosen as you fit the sample to investigate the aims of the study. We will aim to recruit 120 participants to take part in the study.

5. **Do I have to take part?**

You can decide whether to take part in this study. A copy of the information provided here is yours to keep if you do decide to take part. You can still decide to withdraw* during the process of the study without any negative consequences. You do not have to give a reason for your withdrawal. If you wish to withdraw from the research, please contact the lead researcher Bejal Fatania at bfatania1@sheffield.ac.uk.

*Please note that there is a point at which it will not be possible for your data to be withdrawn from the research. This will be upon the submission of your responses as data is collected anonymously and following submission, it is not identifiable for withdrawal.

6. **What will happen to me if I take part? What do I have to do?**

If you wish to take part in the study you will be asked to complete an initial short questionnaire of suspiciousness. If you are selected for the main study you will complete a further two questionnaires of self-esteem and your relationships with others as well as two reasoning tasks.

Following this, you will complete an online task where you will be asked to read a scenario in which you will meet four people and try to figure out how trustworthy or untrustworthy they are.

The duration of the main study from start to finish should approximately take 30 minutes. The collection of this information will support in achieving the research projects objectives and help answer the hypotheses we made.

Upon successful participation of the initial task as well as if selected for the main study, you will be paid for your time and contribution.

7.   What are the possible disadvantages and risks of taking part?

There are no foreseeable disadvantages or risks of taking part in the study.

8.   What are the possible benefits of taking part?

Your participation in the study will be rewarded financially upon successful completion of the study. It is also hoped that this research will contribute to the current understanding of trust and suspiciousness. Moreover, it may provide further information on how practice within mental health services can be adapted to better support clients experiencing varying levels of suspiciousness.

9.   Will my taking part in this project be kept confidential?

All the information that we collect about you during the research will be kept strictly confidential and will only be accessible to members of the research team. Data will be anonymised and therefore you will not be identified in any reports or publications. If you agree to us sharing the information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included.

10.   What is the legal basis for processing my personal data?

According to data protection legislation, we are required to inform you that the legal basis we are applying in order to process your personal data is that 'processing is necessary for the performance of a task carried out in the public interest' (Article 6(1)(e)). Further information can be found in the University's Privacy Notice https://www.sheffield.ac.uk/govern/data-protection/privacy/general.

11.   What will happen to the data collected, and the results of the research project?

Collected data will only be accessed by the research team after it has been anonymised. The anonymised data will be stored on a secure database until the final mark has been received for the research project. The data will then be archived in the University of Sheffield Online Research Data (ORDA).

The results of the research are likely to be disseminated and published. The final project will be uploaded to White Rose eTheses online where you can request a copy of the study and view the published results. You will not be identifiable in any of the publications.

Due to the nature of this research, other researchers may find the collected data useful in answering future research questions. We will ask for your explicit consent for your data to be shared in this way.

12.   Who is organising and funding the research?

No organisation and/or company is funding the research.

13.   Who is the Data Controller?

The University of Sheffield will act as the Data Controller for this study. This means that the University is responsible for looking after your information and using it properly.

14. **Who has ethically reviewed the project?**

This project has been ethically approved via the University of Sheffield's Ethics Review Procedure, as administered by the Psychology department.

15. **What if something goes wrong and I wish to complain about the research?**

If you would like to make a complaint about this project, in the first instance you should contact the lead researcher or their supervisor on the contact details below.

If you feel that your complaint has not been handled to your satisfaction you can contact Professor Elizabeth Milne, Head of the Psychology Department at e.milne@sheffield.ac.uk or Dr Jilly Gibson-Miller, chair of the Department Ethics Subcommittee on jilly.gibson@sheffield.ac.uk.

If your complaint relates to how your personal data has been handled, information about how to raise a complaint can be found in the University's Privacy Notice: https://www.sheffield.ac.uk/govern/data-protection/privacy/general.

16. **Who can I contact for further information?**

Lead Researcher: Bejal Fatania

Email: bfatania1@sheffield.ac.uk

Research Supervisor: Professor Richard Bentall

Email: r.bentall@sheffield.ac.uk

Alternatively, you can email a.sinha@sheffield.ac.uk or leave a telephone message with Amrit Sinha, Research Support Officer on 0114 222 6650 and he will ask the lead researcher to contact you.

I would like to take this opportunity to thank you for your participation in the study.

Appendix M. Consent Form

The University Of Sheffield.

Consent From

A study investigating the rate a person experiencing high and low paranoia learns to trust *Please tick the appropriate boxes*	Yes	No
Taking Part in the Project		
I have read and understood the project information sheet (If you will answer No to this question please do not proceed with this consent form until you are fully aware of what your participation in the project will mean.)	☐	☐
I have been given the opportunity to ask questions about the project.	☐	☐
I agree to take part in the project. I understand that taking part in the project will include completing questionnaires and a computerised task.	☐	☐
I understand that by choosing to participate as a volunteer in this research, this does not create a legally binding agreement nor is it intended to create an employment relationship with the University of Sheffield.	☐	☐
I understand that I can withdraw from the research/study, without notice, at any time point until I click 'submit' at the bottom of the page as submitted survey are not identifiable for removal. I understand that I do not have to give any reasons for why I no longer want to take part and there will be no adverse consequences if I choose to withdraw.	☐	☐
How my information will be used during and after the project		
I understand my personal details such as age and email address etc. will not be revealed to people outside the project.	☐	☐
I understand and agree that other authorised researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I understand and agree that other authorised researchers may use my data in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I give permission for the responses I provide on the questionnaires and the computerised task to be used for future research and learning	☐	☐
So that the information you provide can be used legally by the researchers		
I agree to assign the copyright I hold in any materials generated as part of this project to The University of Sheffield.	☐	☐

Project contact details for further information:

Lead Researcher	Bejal Fatania	Bfatania1@sheffield.ac.uk
Supervisor	Professor Richard Bentall	R.bentall@sheffield.ac.uk
Head of Department	Professor Glenn Waller	G.waller@sheffield.ac.uk

Department of Psychology, Cathedral Court, The University of Sheffield, 1 Vicar Ln, Sheffield City Centre, Sheffield S1 2LT

Submit   No, thank you

Appendix N. Consent Reminder

The University Of Sheffield.

Please read the information below before making your decision

The aim of the current study is to identify whether there is a difference in levels of suspiciousness and the rate a person learns trustworthiness.

You have been selected to participate as you fit the sample to investigate the aims of the study.

Further information about the study and how your information will be used has been provided to you within the project information sheet.

Reminder

- I understand that this study has been ethically approved by the University of Sheffield's Ethics Review Procedure, as administered by the Psychology department.
- I understand that taking part in the study will include completing questionnaires, reasoning tasks, and a computerised task.
- I understand there are no foreseeable disadvantages or risks of taking part in the study.
- I understand that I can withdraw from the study, without notice or a reason, at any time point until I click 'submit' on the survey page at the end of the study.
- I understand that all collected data will be anonymised, kept strictly confidential, and will only be accessible to members of the research team.
- I understand that with my consent the researchers can share the collected information with other researchers, for example, by making it available in a data archive.
- I understand that this research will be used to write a thesis which fulfils part of the doctoral training, and the anonymised results are likely to be disseminated or published.

Continue Participation   **Withdraw**

Appendix O. Study Instructions

Instructions

Please read the instructions below to familiarise yourself with the computerised task

Before you start the Computerised Task

You will be requested to rate your level of trustworthiness for four faces, each appearing individually. Please rate your level of trustworthiness by selecting your rating from the 7-point scale ranging from Very Untrustworthy to Very Trustworthy.
The scale will remain on the screen until you have rated your level of trust for that specific face. Upon selecting the next arrow a new face and scale will appear on the screen. Please rate the level of trustworthiness for all four faces.

Starting the Computerised Task

Once you have completed the initial ratings, you will be prompted to click START and the next button. Please click these buttons when you are ready to complete the computerised task.

During the Computerised Task

A face, a scenario, and a question will be presented on the screen along with a rating scale.
Once you have read the scenario and the question, please rate your level of trust on a 7-point scale ranging from Very Untrustworthy to Very Trustworthy.
The question and the rating scale will remain on the screen until you have responded. Please then select the next arrow.
An outcome for the scenario will then be presented on the screen.
Upon selecting the next arrow, the outcome will disappear, and a new face and scenario will appear on the screen along with a rating scale.
Please take your time to work through the computerised task.

Thank you again for your participation.

Appendix P. Debrief Form

Participant Debrief Form

An experimental study investigating suspiciousness and interpersonal trust

Thank you for participating in this research. Your time and responses are greatly appreciated.

The aim of the current study was to identify whether there is a difference in levels of suspiciousness and the rate a person learns to trust others. You were selected to partake in the study as you fit the sample to investigate the study aims. To assess this, you completed standardised questionnaires, reasoning tasks, and a computerised task. During the study, the researcher did not use any form of deception.

Results of this study will be anonymised and will not include any identifiable information. The study findings are likely to be disseminated and published. You may request a summary of the research findings of this project (once it is completed). To do so, please contact the lead researcher via the email address below.

If you are feeling distressed as a result of your participation, please seek support from the Samaritans on 116 123 (free from any phone), SANEline on 0300 304 7000 (4.30pm–10.30pm every day) or MIND on 0300 123 3393 (9am-6pm, Monday to Friday).

Thank you again for your participation.

Contact Details
Lead Researcher: Bejal Fatania (bfatania1@sheffield.ac.uk)
Supervisor: Professor Richard Bentall (r.bentall@sheffield.ac.uk)
Research Support Officer: Amrit Sinha (a.sinha@sheffield.ac.uk)

Please click the arrow at the bottom of the page to complete the study.

Appendix Q. Ethics Study Approval Letter

Bejal Fatania
Registration number: 190218032
Psychology
Programme: Doctorate in Clinical Psychology

Dear Bejal

PROJECT TITLE: An experimental study investigating suspiciousness and interpersonal trust.
APPLICATION: Reference Number 037290

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 10/02/2021 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 037290 (form submission date: 26/01/2021); (expected project end date: 23/05/2022).
- Participant information sheet 1085468 version 3 (28/05/2021).
- Participant information sheet 1085467 version 3 (28/05/2021).
- Participant information sheet 1086701 version 2 (28/05/2021).
- Participant consent form 1085470 version 3 (14/06/2021).
- Participant consent form 1086703 version 2 (14/06/2021).

The following amendments to this application have been approved:

- Amendment approved: 23/02/2022

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Your responsibilities in delivering this research project are set out at the end of this letter.

Yours sincerely

Department Of Psychology Research Ethics Committee
Ethics Administrator
Psychology

Please note the following responsibilities of the researcher in delivering the research project:
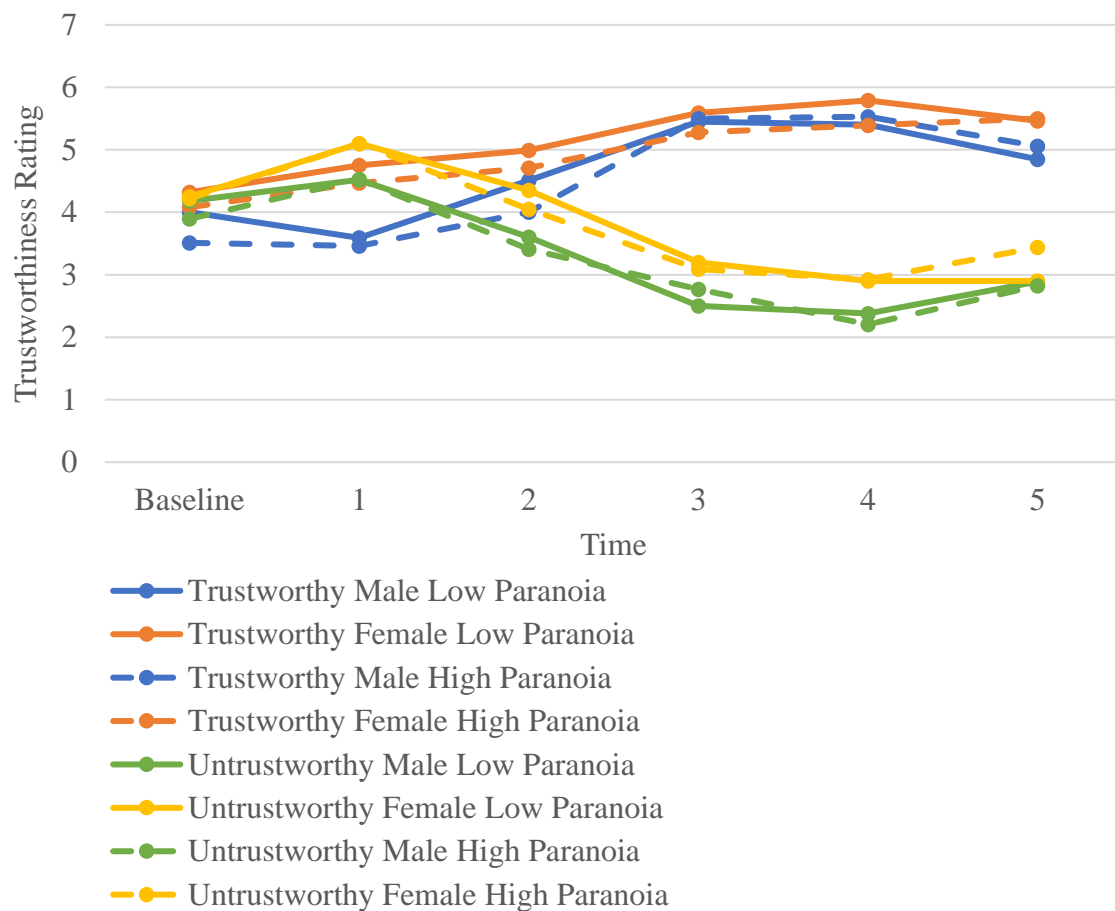
- The project must abide by the University's Research Ethics Policy:
  https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/approval-procedure
- The project must abide by the University's Good Research & Innovation Practices Policy:
  https://www.sheffield.ac.uk/polopoly_fs/1.671066!/file/GRIPPolicy.pdf
- The researcher must inform their supervisor (in the case of a student) or Ethics Administrator (in the case of a member of staff) of any significant changes to the project or the approved documentation.
- The researcher must comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data.
- The researcher is responsible for effectively managing the data collected both during and after the end of the project in line with best practice, and any relevant legislative, regulatory or contractual requirements.

Appendix R. 4-way ANOVA

A 4-way repeated measures ANOVA was also completed. The factors included, the two groups (*low and high paranoia*), the gender of the faces (*male and female*), conditions within the paradigm (trustworthy or untrustworthy) and time (*Baseline-5*)

The 3-way interaction (gender of faces x condition x time) was highly significant $F(5.26, 923.61) = 6.09$, $p < .001$. The 4-way interaction (group x gender of the faces x condition x time) was non-significant, $p = 0.55$.

The findings suggest regardless of being in the low or high paranoia group, trustworthy ratings were higher in the trustworthy condition and lower in the untrustworthy condition across time.

Appendix S. Correlations for the Whole Sample

	Variables	1	2	3	4	5	6	7	Baseline T 8	Baseline T 9	Baseline UT 10	Baseline UT 11	Difference Scores T 12	Difference Scores T 13	Difference Scores UT 14	Difference Scores UT 15
	1. GPTS Total	1	-.39**	-.38**	-.18*	-.04	-.10	-.51**	-.21**	-.07	-.15*	.03	-.01	-.13	.00	-.14
	2. SE Total		1	.53**	.26**	.04	-.10	.37**	.25**	.04	.04	-.06	.024	.15*	.04	-.07
	3. RQ Attachment Anxiety			1	.23**	-.04	-.08	.36**	.30**	.03	.03	-.07	-.06	.12	-.04	.08
	4. RQ Attachment Avoidance				1	-.00	.04	.01	.08	-.03	.06	.03	.08	.02	.00	.01
	5. CRT Total					1	.14	-.07	.05	-.07	.02	-.03	.03	-.01	-.12	.05
	6. JTC Draws to Decision (*n* = 165)						1	-.01	.01	.09	.03	.09	-.12	-.01	.01	.12
	7. Age							1	.27**	.09	.09	.03	.08	.10	.07	.16*
Baseline T	8. Male								1	.22**	-.15*	.24**	.18*	.04	-.01	.06
	9. Female									1	.28**	.24**	-.03	-.05	-.01	.02
Baseline UT	10. Male										1	.30**	-.08	.08	-.04	.02
	11. Female											1	-.05	-.08	.01	-.01
Difference Scores T	12. Male												1	.07	.12	-.09
	13. Female													1	-.04	-.05
Difference Scores UT	14. Male														1	.10
	15. Female															1

** *p* < .01, * *p* < .05

Note. T = Trustworthy Condition, UT = Untrustworthy Condition, GPTS = Green's Paranoia Thought Scale, SE = Self-Esteem, RQ = Relationship Questionnaire, CRT = Cognitive Reflection Test, JTC = Jumping to Conclusion