

# The 'paranoia-as-defence' model of persecutory delusions: A systematic review and meta-analysis

Philip Murphy (DClinPsy)<sup>1,2,3\*</sup>

Richard P. Bentall (PhD)<sup>4</sup>

Daniel Freeman (PhD)<sup>5</sup>

Suzanne O'Rourke (PhD)<sup>6</sup>

Paul Hutton (DClinPsy)<sup>1</sup>

<sup>1</sup>School of Health and Social Care,  
Edinburgh Napier University, Edinburgh, UK

<sup>2</sup>Division of Psychology and Mental Health,  
School of Health Sciences,  
University of Manchester, Manchester, UK

<sup>3</sup>Greater Manchester Mental Health NHS Foundation Trust,  
Prestwich, Manchester, UK

<sup>4</sup>Department of Psychology,  
University of Sheffield, Sheffield, UK

<sup>5</sup>Department of Psychiatry,  
University of Oxford, Oxford, UK

<sup>6</sup>Department of Clinical Psychology,  
School of Health in Social Science,  
University of Edinburgh, Edinburgh, UK

*Correspondence to:*

Dr Philip Murphy  
School of Health and Social Care  
Edinburgh Napier University  
9 Sighthill Court  
Edinburgh  
EH11 4BN

pmurphy1@tcd.ie

## Abstract

### Background

An influential psychological model of persecutory delusions proposed they are caused by a bias towards holding others responsible for negative events, which serves to prevent underlying low self-esteem from reaching awareness. An early (1994) version of the model predicted self-esteem would therefore be preserved in people with these delusions, whereas a later (2001) version suggested it would be unstable, and that there would be a discrepancy between their explicit and implicit self-esteem, with the latter being lower. Our aim was to perform the first comprehensive meta-analytical test of the key predictions of this model, taking into account evidence quality.

### Methods

Reports identified in previous systematic reviews were collated. Electronic databases (i.e., PsycINFO, MEDLINE, EMBASE and Web of Science) were searched from 2012 to September 2016. The review was pre-registered (PROSPERO registration number: CRD42016032782). Cross-sectional data from case-control, longitudinal or experimental studies that examined self-esteem or the externalising attributional bias in individuals diagnosed with schizophrenia-spectrum disorder were eligible for meta-analyses of group differences, but only if at least 50% of participants with psychosis also had current persecutory delusions. Uncontrolled and longitudinal studies were included in meta-analyses of correlations and self-esteem instability, respectively. Study and outcome quality were assessed using the Agency for Healthcare Research and Quality (AHRQ) assessment tool, and a modified version of Grading of Recommendations Assessment, Development and Evaluation (GRADE), respectively.

**Results** We screened 3053 records, examined 104 full-text reports, and included 64 eligible studies. Consistent with the predictions of both versions of the model, paranoia severity in psychosis was positively correlated with the degree of externalising attributional bias ( $k=21$ ,  $N=1128$ ,  $r=0.18$ , 95% CI 0.08, 0.27; moderate quality). People with persecutory delusions also had a greater externalising attributional bias compared to non-clinical ( $k=27$ ,  $N=1442$ ,  $g=0.48$ , 95% CI 0.23, 0.73; moderate quality) and depressed individuals ( $k=10$ ,  $N=421$ ,  $g=1.06$ , 95% CI 0.48, 1.63; moderate quality), and those with psychosis without persecutory delusions ( $k=11$ ;  $N=480$ ;  $g=0.40$ , 95% CI 0.12, 0.68; moderate quality). Contrary to the 1994 version's predictions, paranoia severity in psychosis was negatively correlated with explicit self-esteem ( $k=23$ ,  $N=1866$ ,  $r=-0.26$ , 95% CI -0.34, -0.17; high quality). People with persecutory delusions also had lower explicit self-esteem than non-clinical individuals ( $k=22$ ,  $N=1256$ ,  $g=-0.88$ , 95% CI -1.10, -0.66; high quality) and similarly low explicit self-esteem to people with psychosis without persecutory delusions ( $k=11$ ;  $N=644$ ,  $g=-0.26$ ; 95% CI -0.54, 0.02; moderate quality). Consistent with the 2001 version's predictions, self-esteem instability was positively correlated with paranoia severity in psychosis ( $k=4$ ,  $N=508$ ,  $r=0.23$ , 95% CI 0.11, 0.34; high quality), and people with persecutory delusions had a greater discrepancy between their implicit and explicit self-esteem than depressed individuals ( $k=7$ ,  $N=398$ ,  $g=0.61$ , 95% CI 0.37, 0.85; moderate quality). They had higher explicit self-esteem than depressed individuals ( $k=13$ ,  $N=647$ ,  $g=0.89$ , 95% CI 0.51, 1.28; moderate quality), but similarly low implicit self-esteem ( $k=7$ ,  $N=398$ ,  $g=-0.19$ , 95% CI -0.45, 0.07; very low quality). Contrary to this later version's predictions, they did not have a greater self-esteem discrepancy than non-clinical individuals ( $k=10$ ;  $N=592$ ;  $g=-0.17$ ; 95% CI -0.45 to 0.12; very low quality). There were also no differences between people with psychosis with or without persecutory delusions in implicit self-esteem ( $k=4$ ;  $N=167$ ;  $g=-0.24$ , 95% CI -0.77, 0.30; low quality) or self-esteem discrepancies ( $k=4$ ;  $N=165$ ;  $g=0.17$ , 95% CI -0.19, 0.53; moderate quality).

**Interpretation** Our meta-analytical appraisal of 25 years of research found that, as predicted by the most recent version of the defensive model of persecutory delusions, people with these delusions show an externalising attributional bias, that this and their self-esteem instability are associated with greater paranoia severity, and that they have a greater implicit-explicit self-esteem discrepancy than people with depression. Its predictions that they would have a greater self-esteem discrepancy than non-clinical individuals and people with psychosis without persecutory delusions were not supported, nor was the prediction of the earlier version of the model, that self-esteem would be preserved in persecutory delusions. To overcome the limitations of the observational data we reviewed, experimental studies, which may at some stage include interventionist-causal trials, are now required.

**Funding** None

## **Research in context**

### **Evidence before this study**

We searched PubMed from Jan 1, 1994 to July 31, 2018 for systematic reviews, with or without meta-analyses, which evaluated the defensive model's predictions in relation to persecutory delusions, using the search terms (“attribution\*” OR “externalis\*” OR “personalis\*” OR “self-serving\*” OR “self-esteem” OR “self-worth” OR “self-concept” OR “schema\*”) AND (“psychos\*” OR “schizo\*” OR “delu\*” OR “paranoi\*” OR “persecut\*”). We then reviewed all papers citing the two papers introducing each version of the model, and searched the PROSPERO systematic review database, using keywords “paranoia” and “persecutory delusions”. Three systematic reviews were identified, all of which used narrative synthesis to interpret the evidence. None found clear evidence to support either version of the model, and argued there was evidence against them, however all acknowledged that many of the individual studies were small and therefore lacked power to provide precise estimates, or detect theoretically or clinically relevant findings. For the meta-analyses, we collated all studies included in these reviews. Electronic databases (i.e., PsycINFO, MEDLINE, EMBASE and Web of Science) were then searched from 2012 to September 2016, using the search terms above. The reference lists of all included full-text articles were hand-searched. Relevant authors were contacted where usable but unpublished data were thought to exist.

### **Added value of this study**

To our knowledge, this is the first meta-analysis of 25 years of research testing the key predictions of the defensive account of persecutory delusions. As predicted by both versions of this model, we found moderate quality evidence that persecutory delusions are associated with an externalising attributional bias. Contrary to the early (1994) version's predictions, we found moderate to high quality evidence that people with persecutory delusions have abnormally low explicit self-esteem, and that this is associated with greater paranoia severity. As predicted by the later (2001) version, we found high quality evidence that paranoia severity is associated with greater self-esteem instability, and mixed quality evidence that, compared to people with depression, those with persecutory delusions have greater explicit self-esteem, similarly low implicit self-esteem, and a greater discrepancy between their implicit and explicit self-esteem. However, contrary to the 2001 version, we found very low quality evidence that they have a normal, rather than exaggerated, discrepancy in implicit-explicit self-esteem. Comparisons between people with psychosis with and without current persecutory delusions indicate the former have a heightened externalising attributional bias, but group differences in explicit, implicit and discrepant self-esteem were not evident (low to moderate quality evidence), thus challenging the notion that self-esteem disturbance is specifically associated with these delusions.

### **Implications of all the available evidence**

The claim that persecutory delusions involve defensive processes to protect self-esteem has been influential yet heavily criticised, and a non-defensive account of these delusions has been developed. Unlike previous narrative reviews, our meta-analysis found evidence to support some of the predictions of the later version of this model, but not others. However, the observational research we have reviewed does not allow easy causal inference. Experimental testing of the model is therefore required, whereby the effect on paranoia and persecutory delusions of selectively modifying disputed aspects of the model such as implicit self-esteem is carefully examined. This work, which may at some stage include randomised controlled ‘interventionist-causal’ trials, will require the development of strategies capable of changing

these variables in a way that bypasses explicit self-esteem, as well as more reliable methods of assessing change in implicit self-esteem.

## Introduction

Persecutory (paranoid) delusions involve unfounded beliefs that others are trying to harm the self,(1) and are a major psychiatric problem. They are present in over 70% of patients presenting with a first episode of psychosis,(2) often result in psychiatric hospital admission,(3) and are linked to increased risk of violence.(2) One influential psychological model of these delusions, known as the ‘paranoia as defence’ model,(4,5) proposes they emerge as a consequence of a bias towards holding others responsible for negative events (an externalising attributional bias), in order to reduce awareness of low self-esteem, with low self-esteem conceptualised in the early (1994) version of the model(5) as a discrepancy between one’s actual self and one’s ideal self (Figure 1). It proposed that holding others responsible for negative events is counterproductive, in so far as this may activate fears that others judge the person negatively. This may increase one’s attention to interpersonal threat, which may prompt even more extreme external-personal attributions. Thus, the 1994 version predicts that people with persecutory delusions ought to have a heightened externalising attributional bias and relatively preserved self-esteem, both of which should be related to increased paranoia severity.

The later (2001) version of the model, known as the Attribution–Self-Representation Cycle model,(4) explicitly casts the defensive account within dynamic systems theory, and incorporates social psychological evidence that self-esteem and attributional processes influence each other in a cyclic process as the individual attempts to explain life events. This later version suggests that the externalising attributional bias in the context of persecutory delusions does not provide a complete defence against low self-esteem reaching conscious awareness, and that a combination of the externalising attributional bias and low implicit self-esteem will cause self-esteem to be inherently unstable in relation to greater persecutory delusion severity. It also predicts that covert measurements of self-esteem and attributions will reveal a more negative and self-blaming cognitive architecture than overt assessments, since the former, but not the latter, ought to minimise activation of defensive processes. Thus, two further predictions of the 2001 version are that people with persecutory delusions will have low implicit self-esteem, measured by reaction time or similar tasks, at a level similar to those with depression, and there will be a discrepancy between their implicit and self-reported ‘explicit’ self-esteem, with the latter being higher. It also considers the origins of external-personal causal inferences, drawing on research suggesting they involve less cognitive effort than benign ‘external-situational’ attributions, particularly if an individual has an attentional bias towards threat and difficulty understanding the intentions of others. The early and later versions of this model are depicted in Figures 1 and 2, respectively.

In 2002, Freeman and colleagues proposed an alternative ‘non-defensive’ account of the development and maintenance of persecutory delusions (Figure 3).(6) In this model, persecutory delusions are viewed as threat beliefs, developed in the context of genetic and environmental risk, and maintained by several psychological processes including excessive worry, low self-confidence, intolerance of anxious affect and other internal anomalous experiences, reasoning biases, and the use of safety-seeking strategies.(7) Negative self-beliefs, often developed in the context of adverse inter-personal experiences, mean that the individual feels inferior to others, different and apart, and hence vulnerable. Paranoia feeds on this vulnerability. This model does not predict a discrepancy between implicit and explicit self-esteem, nor does it claim self-esteem instability is central to persecutory delusion development (though it will be common in patients with emotional disorders). The authors regard low self-esteem and negative cognitions as one of a number of interacting causes of persecutory delusions, which are best conceived of as *‘insufficient but non-redundant parts of an unnecessary but sufficient causal condition’*.(7–9) Thus, this model predicts that low self-

esteem is not sufficient for persecutory delusions to form, but may be an essential component of one or more complex pathways. These pathways may be sufficient to cause persecutory delusions, but are not essential, in so far as other complex pathways may also give rise to them.

Given low explicit self-esteem is thought to be common in persecutory delusions, the existence of a defensive causal pathway has proven to be contentious, and the models proposing it have been criticised for lacking parsimony,(9) or being difficult to operationalise.(8) Proponents of the 2001 version of the model place weight on the hypotheses that persecutory delusions involve (1) heightened external-personal attributions, (2) discrepancies between implicit and explicit self-esteem ,(3) low implicit self-esteem, comparable to those with depression, and (4) self-esteem instability, with greater weight currently being placed on the latter.(4) Critics of this, however, place particular weight on the second hypothesis, arguing that even if an external-personal attributional bias is present, its function is moot.(8) Although the 1994 version of the model predicted self-esteem is preserved through the process of making external-personal attributions, there is disagreement over whether the 2001 version makes the same claim. Proponents have argued that the dynamic nature of this later version precludes such predictions, whereas critics have suggested that “*relative preservation of mood and (explicit) self-esteem might be expected*” if this is true.(7–9)

Three recent systematic reviews did not find clear evidence to support either version of this model and, indeed, argue that there is evidence against them.(8,10,11) Each found evidence of low explicit self-esteem but limited or no evidence of an implicit-explicit self-esteem discrepancy in persecutory delusions. Although two found support for an association between persecutory delusions and self-esteem instability,(10,11) one found only mixed evidence that people with persecutory delusions had an exaggerated externalising attributional bias.(8) These reviews had two notable limitations. First, many of the studies in this field are small and therefore unable to reliably detect the full range of important relationships that might in fact exist,(12) suggesting meta-analysis may be required before firm conclusions can be drawn.(13) Second, their assessment of discrepancies between implicit and explicit self-esteem were based on a comparison of the results between groups for each type of self-esteem separately, with just two exceptions.(14,15) However, it has been argued that to adequately test the hypothesis of discrepancy, it is necessary to analyse the difference between implicit and explicit self-esteem within each group as well as differences between groups.(14,15)

The aim of this study was therefore to conduct a systematic review and series of meta-analyses to test key predictions of both the early and later versions of the model. These, and our methodology, were agreed in advance by both one of its leading proponents (RPB) and one of the leading proponents of the non-defensive model (DF). To test the 1994 version, we examined whether people with persecutory delusions have greater explicit self-esteem compared to people with depression and those with psychosis without persecutory delusions, and we also examined whether they have either similar or greater explicit self-esteem compared to non-clinical individuals. We also tested whether paranoia severity in psychosis is positively correlated with explicit self-esteem. To test the predictions of the 2001 version, we examined whether people with persecutory delusions have a greater externalising attributional bias (also predicted by the 1994 version) and greater discrepancy between implicit and explicit self-esteem when compared to non-clinical individuals, people with depression and people with psychosis without persecutory delusions. We also tested the hypothesis that people with persecutory delusions would have similar implicit self-esteem compared to those with depression, but lower implicit self-esteem compared to non-clinical individuals and people with psychosis without persecutory delusions. Moreover, it was predicted that the degree of externalising attributional bias, implicit-explicit self-esteem discrepancy, and self-esteem instability would each be

positively correlated with paranoia severity in psychosis. In contrast, we predicted there would be a negative correlation between implicit self-esteem and paranoia severity in psychosis. Finally, a number of pre-specified moderator analyses were planned to examine the effect of depression and study quality variables on the overall estimates.

## **Methods**

This study adhered to the statement of Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)(16) (see supplement).

### *Search Strategy*

The reports identified in the three systematic reviews published in 2013 and 2014 were firstly collated.(8,10,11) Electronic databases (i.e., PsycINFO, MEDLINE, EMBASE and Web of Science) were then searched by PM (in consultation with PH as well as a research librarian) from 2012 to September 2016. Search terms related to psychosis, delusions, externalising attributional bias and self-esteem were used. The reference lists of all included full-text articles were searched to identify any studies missed in the initial search. Where useable but unpublished data were thought to exist the relevant authors were contacted, and all corresponding authors of included studies were contacted for any further unpublished data (see supplement for full search strategy).

### *Study Selection*

Studies were included in the group comparison analyses if they measured externalising attributional bias, implicit self-esteem or explicit self-esteem in (1) people diagnosed with a schizophrenia spectrum condition (hereafter referred to as “psychosis”) of whom  $\geq 50\%$  had current persecutory delusions and (2) people with depression or non-clinical individuals. Studies comparing people with current persecutory delusions to people with psychosis without persecutory delusions were included in the group comparison analyses providing  $< 50\%$  of the latter group had current persecutory delusions (and, if specified, grandiose delusions). Studies without control group data were eligible for inclusion in the correlation analyses if (1)  $\geq 50\%$  of the sample had psychosis and (2) correlation or regression data was reported between a measure of paranoia/persecutory ideation and a measure of externalising attributional bias or self-esteem. Studies comparing people with current persecutory delusions to people with psychosis without persecutory delusions (irrespective of the presence of grandiose delusions in the latter) were included in the correlation analyses. Cross-sectional data, including baseline data from longitudinal studies, experimental manipulation studies and trials of interventions, were included in the different analyses, with the exception of the self-esteem instability analysis where only longitudinal data were used.

Studies were excluded where  $\geq 50\%$  of the psychosis samples had bipolar disorder, learning disability, a primary diagnosis of substance-induced psychosis or psychosis secondary to a general medical condition. When samples overlapped by  $\geq 25\%$  the study that reported on the largest number of participants was used. Only English-language studies were considered. Selection of studies was conducted by PM in consultation with PH.

### *Outcome Measures and Data Extraction*

Different outcomes were selected corresponding to the different domains of the paranoia as defence model. A ‘data extraction hierarchy’ was developed for most outcomes; this specified what data were most preferable, and what data would be used if these could not be acquired. Because a variety of scoring methods have been proposed for attributional measures, for the externalising attributional bias, the hierarchy was: (1) the external-personal attribution score for



negative events (a measure of the tendency to attribute negative events to other people – rather than to oneself or situational factors); (2) the personalising bias score (a measure of the tendency to attribute negative events to other people rather than to situational factors); (3) the internality attribution score for negative events (a measure of the tendency to attribute negative events to oneself – rather than to other people or situational factors); (4) the externalising bias score (a measure of the tendency to attribute negative, as opposed to positive events, to external causes – either to other people or situational factors). Data from the Internal, Personal, and Situational Attributions Questionnaire (IPSAQ),(17) which can be used to calculate all four indices in the hierarchy above, were preferred over data from the Attributional Style Questionnaire (ASQ),(18) which can only be used to calculate the bottom two indices. Participants' self-ratings were prioritised over independent judges' ratings.

For explicit self-esteem, we preferred to use Rosenberg Self-Esteem Scale (RSES)(19) data; if unavailable, a conceptually equivalent variant was used. We prioritised negative explicit self-esteem data over positive explicit self-esteem data if usable total explicit self-esteem data were unavailable. For implicit self-esteem, we used the following hierarchy: (1) Implicit Association Task (IAT);(20) (2) Emotional Stroop Task (EST);(21,22) (3) Go/No-go Association Task (GNAT).(23) If these data were not available, a conceptually equivalent variant was used. Implicit and explicit self-esteem discrepancies were calculated from the choice of implicit and explicit self-esteem indices above using a statistical method that allowed for the analysis of both within and between group differences, unless already reported (see supplement for details). Self-esteem instability was assessed by the Experience Sampling Method (ESM);(24) or the repeated application of a self-esteem measure such as the RSES (see supplement).

Data were extracted into a spreadsheet by PM and cross-checked by PH. Means and associated standard deviations (SDs) were used for analyses of group differences. Missing SDs were, where possible, calculated from *t* test values, *P*-values, *F*-values, standard errors (SEs) or confidence intervals (CIs) using equations outlined in the Cochrane Handbook(25) and elsewhere.(26) Alternatively, missing SDs were estimated from the mean SD of the other included studies.(27) For within-group analyses, correlation coefficients were extracted, or derived from regression coefficients(28,29) or from group differences between people with psychosis with and without current persecutory delusions using the Campbell Collaboration effect size calculator.(30)

### *Meta-Analytic Calculations*

Meta-analyses were conducted using MetaXL software.(31) For group difference meta-analyses, Hedges' *g* standardised mean differences (SMD) and 95% CIs were computed. When a study had two or more similar groups, these were combined into one using equations specified in the Cochrane Handbook (Version 5.1: Section 7.7.3.8).(25) For correlational meta-analyses, Pearson's correlations were converted into Fisher's *Z* and 95% CIs, and any Spearman correlations were firstly converted into Pearson's correlations.(32) The meta-analytical estimates were then back-transformed into Pearson's correlations to allow interpretation. Following Cohen's conventions,(33) a Hedges' *g* of 0.2, 0.5, and 0.8 were interpreted as small, moderate, and large group differences, respectively, and a Pearson's *r* of 0.1, 0.3, and 0.5 were interpreted as small, moderate, and large correlations, respectively.

Random-effects meta-analyses using the DerSimonian and Laird method (1986)(34) were conducted for all outcomes.(33) When there was less than moderate heterogeneity (i.e.,  $I^2 < 40\%$ ),(25) a sensitivity analysis using a fixed effect analysis was conducted,(35) but reported only if substantively different. Publication bias was assessed through the Doi plot and LFK index for outcomes with at least 10 studies,(25) as this is considered to be a more sensitive

method than the funnel-plot method.(36) However, funnel-plots (see supplement) and Egger's test were also conducted as additional analyses, but the latter was only reported where it differed from the LFK value. The 'trim and fill' method was applied if the LFK index indicated bias ( $LFK > 2$ ). (37)

#### *Moderator Analyses*

Two pre-specified moderators of effect size were examined: (1) matching of groups on demographics [age, gender, education (or a measure of IQ if not reported), ethnicity]; (2) group differences in depression (see supplement for definitions). Random effects meta-regression was used to test these moderator effects using Comprehensive Meta-Analysis version 3 software, but not when fewer than 10 studies in a meta-analysis provided usable data.(25) Two moderator analyses were abandoned due to insufficient data (<5 studies per level of variable): the blinding of the outcome assessor and early vs chronic psychosis.

#### *Subgroup Analysis*

Where group differences in depression significantly moderated an effect size, a subgroup analysis was also conducted to further explore the influence of depression on the relevant effect size. For this analysis, the groups of people with persecutory delusions were coded as either depressed ( $\geq$  mild depression) or non-depressed ( $<$  mild depression) based on a cut-off score on a reported measure of depression (see supplement), and a mixed effects analysis using Comprehensive Meta-Analysis version 3 software was performed.

#### *Risk of Bias and Study Quality*

Following previous reviews,(38) the methodological quality of all studies was assessed using an adapted version of the Agency for Healthcare Research and Quality (AHRQ) assessment tool(39) (see supplement). The quality of the meta-analytical outcomes was assessed using an adapted version of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach(40) (see supplement). The overall GRADE rating (high, moderate, low or very low quality) incorporated considerations of the quality of the studies, publication bias, inconsistency and imprecision.

#### *Registration of Review Protocol and Subsequent Modifications*

The review protocol was reviewed and approved by experts in the psychology of persecutory delusions (DF, RPB) and pre-registered with PROSPERO International Prospective Register of Systematic Reviews (registration number: CRD42016032782). All changes to the protocol were decided before analyses were undertaken, and are fully detailed in the supplement.

## **Results**

Sixty-four studies were identified (Figure 4, Table 1), of which 33 tested hypotheses on the externalising attributional bias, 36 on explicit self-esteem, 11 on implicit self-esteem, 10 on self-esteem discrepancies, and 4 on self-esteem instability. Excluded studies, reasons for exclusion, and a description of included studies are detailed in the supplement. Unpublished data from 6 studies were obtained.(41–46) Studies were published between 1991 and 2016; half ( $k = 32$ ) were conducted in the United Kingdom, with the remainder occurring in Europe ( $k = 17$ ), United States and Canada ( $k = 10$ ) and Australia ( $k = 5$ ).

#### *Study Quality (Table 2, and supplement)*

Consistent methodological problems were non-reporting of pre-specified power calculations and non-blinding of researchers to diagnosis. A lesser problem was a lack of matching of groups on key demographic variables. Most studies selected participants in a relatively unbiased way

(although convenience samples were widely employed), provided adequate sample characteristics, and used valid and reliable measures of diagnostic status, persecutory delusion severity and self-esteem. However, just over a third of the externalising attributional bias measures were judged to be only partially reliable and valid, primarily because they represented the bottom two data extraction hierarchy indices (i.e., they did not distinguish between external-personal and external-situational attributions).

#### *Meta-Analytic Outcomes (Table 2, Figures 5-7, and supplement)*

##### *Externalising Attributional Bias*

As predicted by both versions of the model, people with persecutory delusions had a greater externalising attributional bias than non-clinical individuals ( $k = 27$ ;  $N = 1442$ ;  $g = 0.48$ ; 95% CI = 0.23 to 0.73;  $I^2 = 80\%$ ; moderate quality evidence) and people with psychosis without persecutory delusions ( $k = 11$ ;  $N = 480$ ;  $g = 0.40$ ; 95% CI = 0.12 to 0.68;  $I^2 = 53\%$ ; moderate quality evidence) (see Figure 5). A large difference in externalising between people with persecutory delusions and people with depression was observed, with the former having an exaggerated bias ( $k = 10$ ;  $N = 421$ ;  $g = 1.06$ ; 95% CI = 0.48 to 1.63;  $I^2 = 86\%$ ; moderate quality evidence). There was a small-moderate positive correlation between paranoia severity and the externalising attributional bias in psychosis ( $k = 21$ ;  $N = 1128$ ;  $r = 0.18$ ; 95% CI = 0.08 to 0.27;  $I^2 = 58\%$ ; moderate quality evidence).

##### *Explicit Self-Esteem*

Consistent with the 1994 version, people with persecutory delusions had significantly greater explicit self-esteem than people with depression ( $k = 13$ ;  $N = 647$ ;  $g = 0.89$ ; 95% CI = 0.51 to 1.28;  $I^2 = 80\%$ ; moderate quality evidence). However, contrary to its predictions, people with persecutory delusions had significantly lower explicit self-esteem compared to non-clinical individuals ( $k = 22$ ;  $N = 1256$ ;  $g = -0.88$ ; 95% CI = -1.10 to -0.66;  $I^2 = 68\%$ ; high quality evidence) and similar explicit self-esteem to people with psychosis without persecutory delusions ( $k = 11$ ;  $N = 644$ ;  $g = -0.26$ ; 95% CI = -0.54 to 0.02;  $I^2 = 58\%$ ; moderate quality evidence). Also contrary to the 1994 version, a small-moderate negative correlation between paranoia severity and explicit self-esteem in psychosis was observed ( $k = 23$ ;  $N = 1866$ ;  $r = -0.26$ ; 95% CI = -0.34 to -0.17;  $I^2 = 74\%$ ; high quality evidence).

##### *Implicit Self-Esteem*

Consistent with the predictions of the 2001 version, people with persecutory delusions had lower implicit self-esteem than non-clinical individuals ( $k = 11$ ;  $N = 683$ ;  $g = -0.37$ ; 95% CI = -0.65 to -0.08;  $I^2 = 66\%$ ; low quality evidence) and similar implicit self-esteem to people with depression ( $k = 7$ ;  $N = 398$ ;  $g = -0.19$ ; 95% CI = -0.45 to 0.07;  $I^2 = 34\%$ ; very low quality evidence). However, inconsistent with this version, no significant difference in implicit self-esteem between people with psychosis with and without persecutory delusions was observed ( $k = 4$ ;  $N = 167$ ;  $g = -0.24$ ; 95% CI = -0.77 to 0.30;  $I^2 = 61\%$ ; low quality evidence), nor was there a significant correlation between paranoia severity and implicit self-esteem in psychosis ( $k = 4$ ;  $N = 167$ ;  $r = -0.13$ ; 95% CI = -0.38 to 0.15;  $I^2 = 62\%$ ; low quality evidence).

##### *Discrepancy between Implicit and Explicit Self-Esteem*

As predicted by the 2001 version, people with persecutory delusions had a significantly greater discrepancy between their implicit and explicit self-esteem than people with depression ( $k = 7$ ;  $N = 398$ ;  $g = 0.61$ ; 95% CI = 0.37 to 0.85;  $I^2 = 22\%$ ; moderate quality evidence) (see Figure 6). However, inconsistent with this version, there was no evidence that people with persecutory delusions had a greater implicit-explicit self-esteem discrepancy than non-clinical individuals ( $k = 10$ ;  $N = 592$ ;  $g = -0.17$ ; 95% CI = -0.45 to 0.12;  $I^2 = 61\%$ ; very low quality evidence) or

people with psychosis without persecutory delusions ( $k = 4$ ;  $N = 165$ ;  $g = 0.17$ ; 95% CI = -0.19 to 0.53;  $I^2 = 20\%$ ; moderate quality evidence), and no significant correlation was found between paranoia severity and discrepancy scores in psychosis ( $k = 4$ ;  $N = 165$ ;  $r = 0.09$ ; 95% CI = -0.09 to 0.26;  $I^2 = 15\%$ ; moderate quality evidence).

### *Self-Esteem Instability*

As also predicted by the 2001 version, there was a significant positive correlation between paranoia severity and self-esteem instability in psychosis ( $k = 4$ ;  $N = 508$ ;  $r = 0.23$ ; 95% CI = 0.11 to 0.34;  $I^2 = 38\%$ ; high quality evidence) (see Figure 7). Group differences in self-esteem instability were unavailable.

### *Moderator Analyses (Table 2)*

Differences in severity of depression in people with psychosis significantly moderated the effect size for explicit self-esteem (psychosis with persecutory delusions vs non-clinical individuals) ( $Q = 9.42$ ;  $P = 0.002$ ;  $R^2 = 0.49$ ). When people with persecutory delusions were more depressed, they also had lower explicit self-esteem ( $B = -0.70$ ;  $SE = 0.23$ ;  $P = 0.002$ ). However, the test of residual heterogeneity was significant ( $Q = 31.71$ ;  $P = 0.003$ ), suggesting unexplained variance in explicit self-esteem group differences. No other moderator analyses were significant.

### *Subgroup Analysis*

A subgroup analysis on the explicit self-esteem data showed that individuals with persecutory delusions and depression had significantly lower explicit self-esteem than non-clinical individuals ( $k = 12$ ;  $N = 698$ ;  $g = -0.99$ ; 95% CI = -1.28 to -0.70;  $z = -6.71$ ;  $P = <0.001$ ), but that there was no significant difference in explicit self-esteem between those with persecutory delusions who were not depressed and non-clinical individuals ( $k = 5$ ;  $N = 296$ ;  $g = -0.51$ ; 95% CI = -1.09 to 0.08;  $z = -1.69$ ;  $P = 0.091$ ). However, the difference between the 2 effect sizes was not significantly different ( $Q = 2.09$ ;  $P = 0.148$ ).

### *Publication Bias (Table 2)*

Potential publication bias was indicated for the analyses of externalising attributional bias and explicit self-esteem (psychosis with persecutory delusions vs depression) (see supplement for funnel-plots). However, the 'trim and fill' method<sup>35</sup> did not impute any missing studies; thus, the point estimates remained the same.

## **Discussion**

Over the last 25 years, 64 individual studies involving 5363 participants (3562 participants with psychosis, 442 participants with depression, and 1359 non-clinical individuals) have tested the paranoia as defence model of persecutory delusions. To overcome the power limitations of individual studies, we conducted the first meta-analytical appraisal of published and unpublished evidence relating to both the 1994 and 2001 versions of the model. We also employed a method for calculating the discrepancy between implicit and explicit self-esteem, thus enabling the first analysis of both within and between group differences in this outcome, and our pre-registered protocol was approved by exponents of both the general paranoia as defence model(4,5) and an alternative non-defensive model.(6)

Proponents of the defensive model will be encouraged by our finding that people with persecutory delusions do indeed have an elevated externalising attributional bias. This appears to be specific to persecutory delusions, and it is associated with paranoia severity. Proponents of the 2001 version of the model will also be encouraged by the observed association between self-esteem instability and paranoia severity, and our finding that those with persecutory

delusions and those with depression have similarly low implicit self-esteem. That people with persecutory delusions have better explicit self-esteem than those with depression suggests they may indeed have an implicit-explicit self-esteem discrepancy relative to this group, and a direct comparison of their self-esteem discrepancies appeared to confirm this. Although their explicit self-esteem was considerably lower than non-clinical individuals, the moderator analysis suggests this may at least partly be a function of co-occurring depression.

Critics of the model, on the other hand, might reasonably note that an elevated externalising attributional bias does not in itself tell us anything about its function,(6,8,9) and they may also be concerned by the conceptual overlap between holding others responsible for negative events and worrying that others wish to cause one harm.(6,47) In relation to the 2001 version of the model, they may acknowledge that self-esteem instability is linked to paranoia severity, but they may see no need to invoke defensive explanations for this, and query the specificity of its effects to paranoia. They may argue that the low level of implicit self-esteem observed in people with persecutory delusions is also predicted by the non-defensive model,(6) and that defensive accounts are not needed to explain why lower explicit self-esteem is associated with greater paranoia. They may query whether the pattern of self-esteem findings is attributable to the characteristics of people with persecutory delusions, or whether they actually tell us more about the self-esteem profile of people with depression. Moreover, a central concern of critics of the 2001 version may be that its specific claim of an exaggerated self-esteem discrepancy in people with persecutory delusions was not supported by the evidence, and that a non-defensive account(6) reflects a more parsimonious interpretation of the pattern of findings when comparisons with non-clinical individuals were considered. Although explicit self-esteem may be higher when people with persecutory delusions are not depressed, this applies to relatively few people with these difficulties(48,49) – an observation which has been claimed to be inconsistent with a defensive model, or at the least the earlier 1994 version. There was also no difference in implicit or discrepant self-esteem between those with psychosis with and without persecutory delusions, thus casting doubt over claims of specificity. Finally, the heterogeneity in many of the estimates reduces the quality of the conclusions that can be drawn, both for and against the model.

Proponents might counter that the correlation between the externalising attributional bias and paranoia severity observed here was small-moderate in magnitude, which is inconsistent with there being significant conceptual overlap between these variables. They may argue that the 2001 version of the model successfully predicted that greater instability of self-esteem would be associated with greater paranoia severity(4) and, although there may indeed be non-defensive explanations for this relationship, one strength of a model is its ability to make specific predictions that are subsequently supported by evidence. Although it is true that the relationship between explicit self-esteem and paranoia does not require invocation of defensive processes, it does not preclude them either. Indeed, both models predict a reciprocal relationship between paranoia and low explicit self-esteem. In relation to the self-esteem discrepancy findings, one's view on the nature of the discrepancy in non-clinical and depressed individuals is key to interpreting them. If non-clinical individuals do have a self-esteem discrepancy, and depressed individuals do not, then it follows from the meta-analytical findings presented here that a self-esteem discrepancy, albeit not abnormal or exaggerated, also characterises people with persecutory delusions. Such a self-esteem discrepancy, it could be argued, might at the very least be consistent with what critics have referred to as the 'weak' version of the defensive model, which provides for scenarios whereby the externalising attributional bias in the context of persecutory delusions only partially fulfills its defensive function (i.e., it does not fully preserve explicit self-esteem but prevents explicit self-esteem from falling to the even lower level of implicit self-esteem).(50) Although many people with persecutory delusions are indeed

depressed,(48,49) there is also evidence that fluctuations in mood are strongly associated with the formation and maintenance of paranoia in the general population,(51) which would be consistent with the 2001 version's predictions. Moreover, if depression is common in persecutory delusions, then an adequate test of the defensive model requires controlling for this in some way, since depression may indicate that defensive processes are failing to adequately maintain a self-esteem discrepancy. Thus, if people with persecutory delusions are generally depressed, then comparisons with people with depression alone may be better placed to tell us what is specific to persecutory delusions. Indeed, since depression is elevated in people with persecutory delusions, it is unclear why they do not have a much *smaller* discrepancy between their implicit and explicit self-esteem, when compared to non-clinical individuals. Although the current findings might be taken to mitigate against claims of a specific relationship between low self-esteem (whether explicit or implicit) and the presence of persecutory delusions in psychosis, proving specificity is difficult, and it is plausible that people with psychosis without current persecutory delusions nonetheless continue to carry the self-esteem risk factors for them. If the 2001 version of the defensive model represents an "*unnecessary but sufficient*"(52) cause of persecutory delusions, then threats to self-esteem, low implicit self-esteem and the presence of an externalising attributional bias should each be regarded as "*insufficient but non-redundant*"(52) components of this process, therefore all may be required for persecutory delusion occurrence – something few studies have measured. Finally, heterogeneity in meta-analytical estimates is often taken to reflect the presence of unknown moderators, and might therefore be viewed as informative. Indeed, it was variance in cross-sectional estimates of self-esteem in persecutory delusions that first motivated researchers to examine whether self-esteem instability itself might be tied to paranoia severity.(4)

Critics may respond to these counterarguments by noting that the modest correlation between the externalising attributional bias and paranoia severity could be a function of there being only a modest conceptual overlap between the measures of these variables, and that if the defensive account cannot detect differences between people with current persecutory delusions and those with psychosis without persecutory delusions (even those who had such difficulties in the past) then another model would need to account for actual delusion occurrence. They may also argue that people with psychosis without persecutory delusions certainly do provide a better matched control for testing the specificity of a model of paranoia, and that fluctuations in self-esteem, emotional dysregulation, and mood instability are common in many mental health disorders, especially where negative self-views are involved (e.g., depression, eating disorders, and borderline personality disorders), and that these are part-and-parcel of emotional difficulties and not a sign of defence processes. Negative experiences will likely trigger fluctuations in emotional state, and mood instability is also noted to occur in hallucinations(53) and has even been shown to be a vulnerability factor for depression.(54) They may conclude that a strength of a model is not just its predictive power, but also its parsimony – i.e., that all of the findings reported here and elsewhere can be explained without the need to invoke defensive processing.

Overall then, there is clearly much room for continued disagreement. Based on our findings, proponents and critics of the defensive model are, however, likely to now agree that persecutory delusions are associated with an externalising attributional bias. However, whether this bias has a key causal or defensive function is likely to continue to be debated. Although our self-esteem discrepancy findings demonstrate that persecutory delusions are unlikely to involve an exaggerated or abnormal self-esteem discrepancy, proponents of the 2001 version of the model will note that interpreting their overall pattern turns on whether one believes that non-clinical individuals do have a discrepancy, and whether people with depression do not. Complicating resolution of this debate, however, are concerns over the validity of measures of current

measures of implicit self-esteem.(55) Noting this, Buhrmester and colleagues have argued for a different approach to measuring this construct:(55)

*“To circumvent [self-presentational processes], we suggest that respondents be interviewed as they reflect on their self-worth with an eye to illuminating (a) retrospective support for people’s assertions about themselves and (b) potential contradictions between people’s claims about their self-worth and their putative evidence for such claims. Defensiveness shows signs that people possess self-evaluations that they do not “own” when they engage in deliberate self-report.”(55)*

What is also clear from our meta-analysis is that people with persecutory delusions do indeed have abnormally low self-esteem (whether explicit or implicit), when compared to non-clinical individuals. Thus, the 1994 version’s hypothesis, that self-esteem is preserved by the operation of defensive attributional processes, can be rejected. Moreover, there is mounting experimental evidence that selectively improving explicit self-esteem can cause improvements in clinical and non-clinical paranoia.(56–58) Thus it seems clear that, regardless of whether defensive processes are also at play, negative explicit self-esteem is likely to contribute directly to paranoia.

Proponents and critics will also agree that observational research generally precludes firm causal inferences, even with the enhanced power afforded by meta-analysis. Thus, there now needs to be experimental testing of the 2001 version of the model, whereby the effect on paranoia and persecutory delusions of selectively manipulating attributional style, implicit self-esteem and self-esteem instability is carefully examined. This work, which may – if warranted – include randomised controlled ‘interventionist-causal’ trials,(59,60) will require the development of strategies capable of changing these variables without also changing explicit self-esteem, as well as more reliable methods of assessing change in implicit self-esteem. Perhaps the most important test involves measuring the effect of manipulating implicit self-esteem alone on paranoia. If a significant and selective improvement in implicit self-esteem can be achieved, and if this causes an improvement in persecutory delusions and self-esteem stability, then this would be strong support for the model.

### *Limitations*

Resource constraints meant that we were limited to English language studies; however, it is unlikely that data excluded for this reason would have had a substantial impact upon the reported effect sizes or our conclusions. There was also an insufficient number of studies to carry out some of the planned moderator analyses and tests of publication bias. In addition, while most of the analyses produced reasonable quality evidence, the estimates for implicit self-esteem were less reliable, in part because of the methodological problems related to measuring this.(10,11,61) Finally, the relative complexity of this review made it challenging to minimise the time-lag between search completion and publication; however, we are not aware of any major new studies being published since the initial search was completed.

### *Conclusion*

An influential model of persecutory delusions(4,5) proposed they are caused or maintained by a heightened bias towards holding others responsible for negative events, and that this bias helps to prevent low self-esteem from reaching awareness. Our meta-analytical appraisal of 25 years of research provides evidence to support several predictions of the 2001 version of the model, but also some evidence that does not. The 1994 version’s prediction, that persecutory delusions may involve preserved self-esteem, can be rejected. Experimental research that manipulates the key variables of interest, which may at some stage include interventionist-

causal trials,(59) is now required to help resolve the debate, and to determine whether the defensive model has clinical implications for people with persecutory delusions.

### **Declaration of Interests**

PM, PH and SOR report no conflicts of interest. RPB developed the paranoia as defence model of persecutory delusions. DF developed an alternative model, known as the threat anticipation model of persecutory delusions. He has written popular science, self-help, and academic books about paranoia with several publishers for which royalties are received and he is also a co-founder of Oxford VR, a University of Oxford spinout company.

### **Contributors**

PM and PH conceived the study. PM wrote the first draft of the protocol. PH, RPB, and DF substantially contributed to the critical revision of the protocol. PM, PH, RPB, and DF read and approved the final protocol. PM conducted the literature search, selected the studies, extracted the data, analysed the data, assessed study and outcome quality, interpreted the data, and wrote the first draft of the manuscript. PH consulted in the literature search and the selection of the studies, cross-checked the data and the quality assessments, consulted in the data analysis, and substantially contributed to the interpretation of the data and the critical revision of the manuscript. RPB, DF, and SOR substantially contributed to the interpretation of the data and the critical revision of the manuscript. All authors read and approved the final manuscript.

### **References**

1. Freeman D, Garety PA. Comments on the content of persecutory delusions: Does the definition need clarification? *Br J Clin Psychol* [Internet]. 2000 Nov;39(4):407–14. Available from: <http://doi.crossref.org/10.1348/014466500163400>
2. Coid JW, Ullrich S, Kallis C, Keers R, Barker D, Cowden F, et al. The Relationship Between Delusions and Violence. *JAMA Psychiatry* [Internet]. 2013 May 1;70(5):465. Available from: <http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2013.12>
3. Castle DJ, Phelan M, Wessely S, Murray RM. Which Patients with Non-affective Functional Psychosis are not Admitted at First Psychiatric Contact? *Br J Psychiatry* [Internet]. 1994 Jul 2;165(01):101–6. Available from:



- [https://www.cambridge.org/core/product/identifier/S0007125000072263/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S0007125000072263/type/journal_article)
4. Bentall RP, Corcoran R, Howard R, Blackwood N, Kinderman P. PERSECUTORY DELUSIONS: A REVIEW AND THEORETICAL INTEGRATION. *Clin Psychol Rev* [Internet]. 2001 Nov;21(8):1143–92. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0272735801001064>
  5. Bentall RP, Kinderman P, Kaney S. The self, attributional processes and abnormal beliefs: Towards a model of persecutory delusions. *Behav Res Ther* [Internet]. 1994 Mar;32(3):331–41. Available from: <http://linkinghub.elsevier.com/retrieve/pii/0005796794901317>
  6. Freeman D, Garety PA, Kuipers E, Fowler D, Bebbington PE. A cognitive model of persecutory delusions. *Br J Clin Psychol* [Internet]. 2002 Nov;41(4):331–47. Available from: [http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L35439327%5Cnhttp://dx.doi.org/10.1348/014466502760387461%5Cnhttp://sfx.metabib.ch/sfx\\_locator?sid=EMBASE&issn=01446657&id=doi:10.1348/014466502760387461&atitle=A+cognitive+model+of+](http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L35439327%5Cnhttp://dx.doi.org/10.1348/014466502760387461%5Cnhttp://sfx.metabib.ch/sfx_locator?sid=EMBASE&issn=01446657&id=doi:10.1348/014466502760387461&atitle=A+cognitive+model+of+)
  7. Freeman D. Persecutory delusions: a cognitive perspective on understanding and treatment. *The Lancet Psychiatry*. 2016.
  8. Garety PA, Freeman D. The past and future of delusions research: from the inexplicable to the treatable. *Br J Psychiatry* [Internet]. 2013 Nov 2;203(05):327–33. Available from: [https://www.cambridge.org/core/product/identifier/S0007125000053538/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S0007125000053538/type/journal_article)
  9. Freeman D, Garety P. Advances in understanding and treating persecutory delusions: a review. *Soc Psychiatry Psychiatr Epidemiol* [Internet]. 2014 Aug 9;49(8):1179–89. Available from: <http://link.springer.com/10.1007/s00127-014-0928-7>
  10. Kesting M-L, Lincoln TM. The relevance of self-esteem and self-schemas to persecutory delusions: A systematic review. *Compr Psychiatry* [Internet]. 2013 Oct;54(7):766–89. Available from: <http://dx.doi.org/10.1016/j.comppsy.2013.03.002>
  11. Tiernan B, Tracey R, Shannon C. Paranoia and self-concepts in psychosis: A systematic review of the literature. *Psychiatry Res* [Internet]. 2014 May;216(3):303–13. Available from: <http://dx.doi.org/10.1016/j.psychres.2014.02.003>
  12. Maxwell SE. The Persistence of Underpowered Studies in Psychological Research: Causes, Consequences, and Remedies. *Psychol Methods* [Internet]. 2004;9(2):147–63. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/1082-989X.9.2.147>
  13. Borenstein M, Hedges L, Higgins J, Rothstein H. Vote Counting – A New Name for an Old Problem. In: *Introduction to Meta-Analysis*. 2009. p. 251–5.
  14. Kesting M-L, Mehl S, Rief W, Lindenmeyer J, Lincoln TM. When paranoia fails to enhance self-esteem: Explicit and implicit self-esteem and its discrepancy in patients with persecutory delusions compared to depressed and healthy controls. *Psychiatry Res* [Internet]. 2011 Apr;186(2–3):197–202. Available from: <http://dx.doi.org/10.1016/j.psychres.2010.08.036>
  15. Vázquez C, Diez-Alegría C, Hernández-Lloreda MJ, Moreno MN. Implicit and explicit self-schema in active deluded, remitted deluded, and depressed patients. *J Behav Ther Exp Psychiatry* [Internet]. 2008 Dec;39(4):587–99. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0005791608000207>
  16. Moher D, Liberati A, Tetzlaff J, Altman D, Group TP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Ann Intern Med* [Internet]. 2009 Aug 18;151(4):264. Available from: <http://annals.org/article.aspx?doi=10.7326/0003-4819-151-4-200908180-00135>

17. Kinderman P, Bentall R. A new measure of causal locus: The Internal, Personal and Situational Attributions Questionnaire. *Pers Individ Dif*. 1996;20:261–4.
18. Peterson C, Semmel A, von Baeyer C, Abramson LY, Metalsky GI, Seligman MEP. The attributional Style Questionnaire. *Cognit Ther Res* [Internet]. 1982 Sep;6(3):287–99. Available from: <http://link.springer.com/10.1007/BF01173577>
19. Rosenberg M. Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy. Measures package. 1965. 61 p.
20. Greenwald AG, McGhee DE, Schwartz JLK. Measuring individual differences in implicit cognition: The implicit association test. *J Pers Soc Psychol* [Internet]. 1998;74(6):1464–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9654756>
21. Williams JMG, Mathews A, MacLeod C. The emotional Stroop task and psychopathology. *Psychol Bull*. 1996;120(1):3–24.
22. Stroop J. Stroop color word test. *J Exp Psychol*. 1935;18:643–62.
23. Nosek B a, Banaji MR. The Go/No-Go Association Task. *Soc Cogn* [Internet]. 2001 Dec;19(6):625–66. Available from: <http://guilfordjournals.com/doi/10.1521/soco.19.6.625.20886>
24. Csikszentmihalyi M, Larson R. Validity and Reliability of the Experience- Sampling Method. *J Nerv Ment Disord*. 1987;175(9):526–36.
25. Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. 2011.
26. Borenstein M, Hedges L, Higgins J, Rothstein H. Converting Among Effect Sizes. In: *Introduction to Meta-Analysis*. 2009. p. 45–9.
27. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* [Internet]. 2006 Jan;59(1):7–10. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0895435605003227>
28. Peterson RA, Brown SP. On the Use of Beta Coefficients in Meta-Analysis. *J Appl Psychol* [Internet]. 2005;90(1):175–81. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0021-9010.90.1.175>
29. Kelley K, Maxwell SE. Sample Size for Multiple Regression: Obtaining Regression Coefficients That Are Accurate, Not Simply Significant. *Psychol Methods* [Internet]. 2003;8(3):305–21. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/1082-989X.8.3.305>
30. Wilson D. Practical Meta-analysis Effect Size Calculator. <https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php>. 2017.
31. Barendregt J. *MetaXL User Guide version 5.3*. 2016.
32. Rupinski M, Dunlap W. Approximating Pearson Product-Moment Correlations from Kendall's Tau and Spearman's Rho. *Educ Psychol Meas*. 1996;56(3):419–29.
33. Cohen J. *Statistical power analysis for the behavioural sciences*. 2nd. 1988. 567 p.
34. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–88.
35. Borenstein M, Hedges L, Higgins J, Rothstein H. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*. 2010;97–111.
36. Furuya-Kanamori L, Barendregt JJ, Doi SAR. A new improved graphical and quantitative method for detecting bias in meta-analysis. *Int J Evid Based Healthc* [Internet]. 2018 Apr;1. Available from: <http://insights.ovid.com/crossref?an=01787381-900000000-99914>
37. Duval S, Tweedie R. Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics* [Internet]. 2000 Jun;56(2):455–63. Available from: <http://doi.wiley.com/10.1111/j.0006->

- 341X.2000.00455.x
38. Larkin A, Hutton P. Systematic review and meta-analysis of factors that help or hinder treatment decision-making capacity in psychosis. *Br J Psychiatry* [Internet]. 2017 Oct 2;211(04):205–15. Available from: [https://www.cambridge.org/core/product/identifier/S0007125000280392/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S0007125000280392/type/journal_article)
  39. Williams J, Plassman B, Burke J, Benjamin S, Benjamin S. Preventing Alzheimer's disease and cognitive decline. 2010.
  40. Guyatt G, Oxman A, Vist G, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Br Med J* [Internet]. 2008;336(April):1106–10. Available from: [www.uptodate.com](http://www.uptodate.com)
  41. Berry K, Bucci S, Kinderman P, Emsley R, Corcoran R. An investigation of attributional style, theory of mind and executive functioning in acute paranoia and remission. *Psychiatry Res* [Internet]. 2015 Mar;226(1):84–90. Available from: <http://dx.doi.org/10.1016/j.psychres.2014.12.009>
  42. Fornells-Ambrojo M, Garety PA. Understanding attributional biases, emotions and self-esteem in poor me paranoia: Findings from an early psychosis sample. *Br J Clin Psychol* [Internet]. 2009 Jun;48(2):141–62. Available from: <http://doi.crossref.org/10.1348/014466508X377135>
  43. McKay R, Langdon R, Coltheart M. Paranoia, persecutory delusions and attributional biases. *Psychiatry Res* [Internet]. 2005 Sep;136(2–3):233–45. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0165178105001666>
  44. McKay R, Langdon R, Coltheart M. The defensive function of persecutory delusions: An investigation using the Implicit Association Test. *Cogn Neuropsychiatry* [Internet]. 2007 Jan;12(1):1–24. Available from: <http://www.tandfonline.com/doi/abs/10.1080/13546800500363996>
  45. Mehl S, Landsberg MW, Schmidt A-C, Cabanis M, Bechdolf A, Herrlich J, et al. Why Do Bad Things Happen to Me? Attributional Style, Depressed Mood, and Persecutory Delusions in Patients With Schizophrenia. *Schizophr Bull* [Internet]. 2014 Nov 1;40(6):1338–46. Available from: <https://academic.oup.com/schizophreniabulletin/article-lookup/doi/10.1093/schbul/sbu040>
  46. Sundag J, Lincoln TM, Hartmann MM, Moritz S. Is the content of persecutory delusions relevant to self-esteem? *Psychosis* [Internet]. 2015 Jul 3;7(3):237–48. Available from: <http://dx.doi.org/10.1080/17522439.2014.947616>
  47. Freeman D, Garety P. Advances in understanding and treating persecutory delusions: A review. Vol. 49, *Social Psychiatry and Psychiatric Epidemiology*. 2014. p. 1179–89.
  48. Buckley PF, Miller BJ, Lehrer DS, Castle DJ. Psychiatric comorbidities and schizophrenia. *Schizophr Bull*. 2009;35(2):383–402.
  49. Vorontsova N, Garety P, Freeman D. Cognitive factors maintaining persecutory delusions in psychosis: The contribution of depression. *J Abnorm Psychol* [Internet]. 2013;122(4):1121–31. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0034952>
  50. Garety PA, Freeman D. Cognitive approaches to delusions: A critical review of theories and evidence. *Br J Clin Psychol* [Internet]. 1999 Jun;38(2):113–54. Available from: <http://doi.crossref.org/10.1348/014466599162700>
  51. Marwaha S, Broome MR, Bebbington PE, Kuipers E, Freeman D. Mood instability and psychosis: Analyses of British national survey data. *Schizophr Bull*. 2014;
  52. Mackie JL. *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press; 1974.
  53. Marwaha S, Bebbington P. Mood as a mediator of the link between child sexual abuse

- and psychosis. *Soc Psychiatry Psychiatr Epidemiol*. 2015;
54. Franck E, De Raedt R. Self-esteem reconsidered: Unstable self-esteem outperforms level of self-esteem as vulnerability marker for depression. *Behav Res Ther*. 2007;
  55. Buhrmester MD, Blanton H, Swann WB. Implicit Self-Esteem: Nature, Measurement, and a New Way Forward. *J Pers Soc Psychol*. 2011;
  56. Freeman D, Pugh K, Dunn G, Evans N, Sheaves B, Waite F, et al. An early Phase II randomised controlled trial testing the effect on persecutory delusions of using CBT to reduce negative cognitions about the self: The potential benefits of enhancing self confidence. *Schizophr Res*. 2014;160(1–3):186–92.
  57. Freeman D, Evans N, Lister R, Antley A, Dunn G, Slater M. Height, social comparison, and paranoia: An immersive virtual reality experimental study. *Psychiatry Res*. 2014;
  58. Atherton S, Antley A, Evans N, Cernis E, Lister R, Dunn G, et al. Self-Confidence and Paranoia: An Experimental Study Using an Immersive Virtual Reality Social Situation. *Behav Cogn Psychother*. 2016;
  59. Kendler KS, Campbell J. Interventionist causal models in psychiatry: repositioning the mind–body problem. *Psychol Med* [Internet]. 2009 Jun 10;39(06):881. Available from: [http://www.journals.cambridge.org/abstract\\_S0033291708004467](http://www.journals.cambridge.org/abstract_S0033291708004467)
  60. Freeman D, Dunn G, Startup H, Pugh K, Cordwell J, Mander H, et al. Effects of cognitive behaviour therapy for worry on persecutory delusions in patients with psychosis (WIT): A parallel, single-blind, randomised controlled trial with a mediation analysis. *The Lancet Psychiatry*. 2015;2(4):305–13.
  61. Bosson JK, Swann WB, Pennebaker JW. Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *J Pers Soc Psychol* [Internet]. 2000;79(4):631–43. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.79.4.631>

**Figure 1. The ‘paranoia-as-defence’ model: Bentall et al. (1994)**

**Negative events**  
Threatening self-esteem (violating self-standards)



**Actual-ideal self discrepancy**  
(lowered explicit self-esteem)



**Externalising attributional bias**



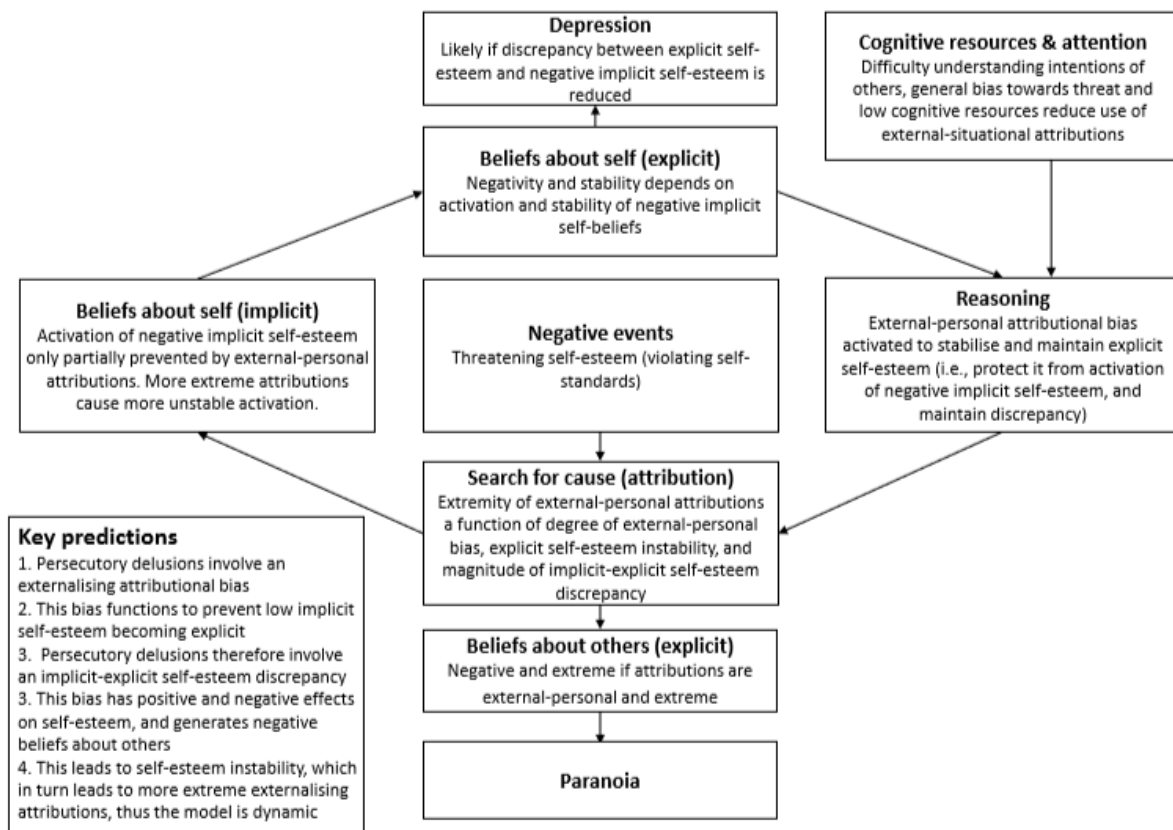
**Reduction in actual-ideal self discrepancy**  
(improved explicit self-esteem)

**Self-other discrepancy**  
(e.g., "other people think badly of me")

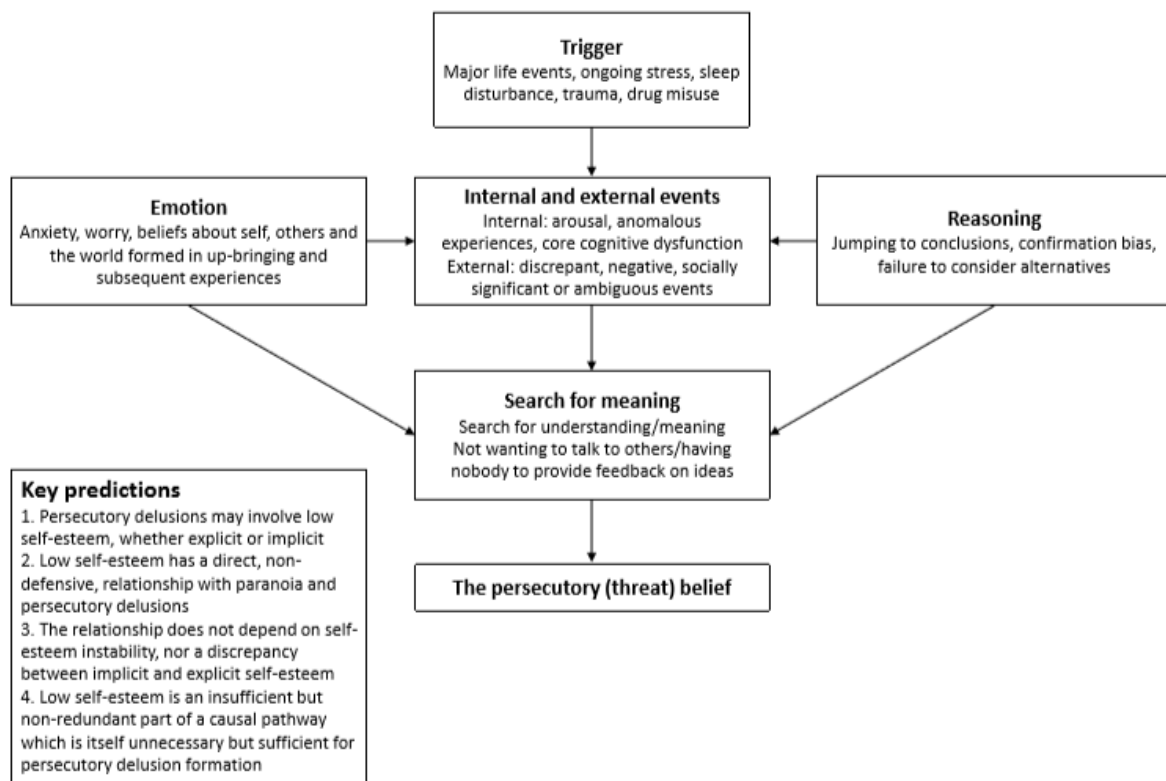
**Key predictions**

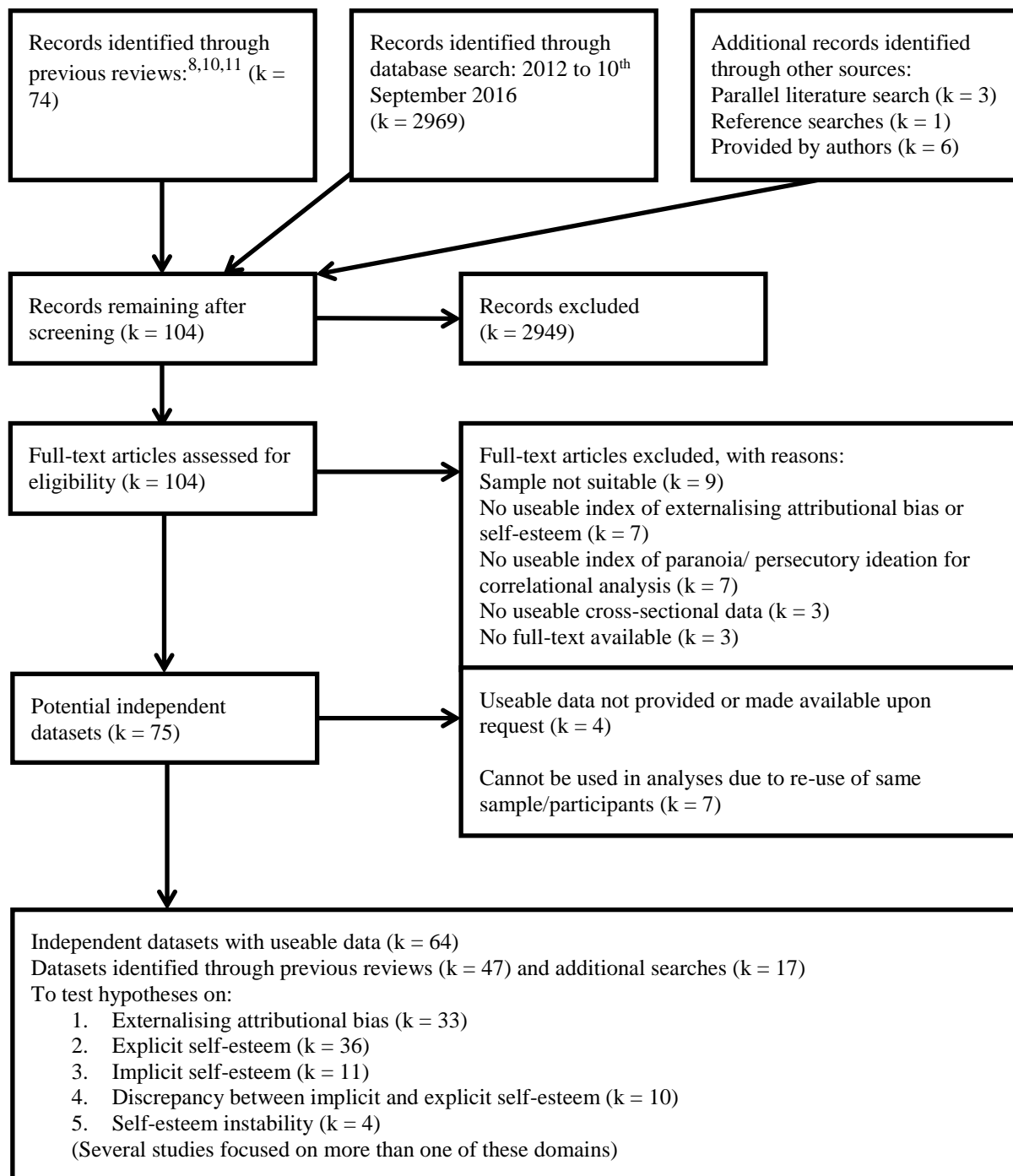
1. Persecutory delusions involve an externalising attributional bias
2. This bias protects self-esteem, but causes an individual to perceive others as thinking negatively of them
3. This self-other discrepancy increases interpersonal threat, which increases the likelihood of external attributions

Figure 2. The revised ‘paranoia-as-defence’ model: Bentall et al. (2001)



**Figure 3: A cognitive model of persecutory delusions: Freeman et al. (2002)**





**Figure 4. PRISMA Flowchart of Study Selection**



**Table 1. Studies Included in the Meta-Analysis<sup>a</sup>**

<b>Study Ref (First Author, Year)</b>	<b>Participant Group/s<sup>b</sup> (N in Parentheses)</b>	<b>Relevant Domain/s</b>
Aakre, 2009	Current PDs (18); Remitted PDs (30); Remitted non-PD delusions (17); Non-clinical (29)	EAB
Bentall, 1991	Current PDs (17); Depression (17); Non-clinical (17)	EAB
Bentall, 2005	Current PDs (16); Depression (16); Non-clinical (16)	EAB
Bentall, 2008	Current PDs (39); Remitted PDs (29); Depression (27); Non-clinical (33)	ESE
Ben-Zeev, 2009	Psychosis (194)	ESE
Berry, 2015 <sup>c</sup>	Current PDs (25); Non-clinical (25)	EAB
Besnier, 2011	Current PDs (30); Non-clinical (60)	ISE
Candido, 1990	Non-depressed PDs (15); Depressed PDs (15); Depression (15)	EAB; ESE
Carlin, 2005	Current PDs (31); Non-PD psychosis (34)	EAB
Collett, 2016	Current PDs (21); Non-clinical (21)	ESE
Combs, 2009	Current PDs (32); Non-PD delusions (28); Non-clinical (50)	EAB; ESE
Diez-Alegria, 2006	Current PDs (40); Remitted PDs (25); Depression (35); Non-clinical (36)	EAB
Erickson, 2012	Psychosis (57)	ESE; SEI
Espinosa, 2014	Current PDs (79); Depression (38); Non-clinical (52)	ESE; ISE; SED
Fear, 1996	Current PDs (20); Non-PD delusions (9); Non-clinical (20)	EAB
Fornells-Ambrojo, 2009 <sup>c</sup>	Current PM PDs (20); Depression (21); Non-clinical (32)	EAB; ESE
Freeman, 1998	Current PDs (28); Non-PD delusions (25);	ESE
Freeman, 2012	Psychosis (130)	ESE
Garety, 2013	Current PDs (118); Current PGDs (52); Non-PGD psychosis (43)	ESE
Humphreys, 2006	Current PDs (15); Non-PD psychosis (20)	EAB; ESE
Janssen, 2006	Psychosis (23)	EAB
Jolley, 2006	Current PDs (7); Current PGDs (7); Non-PD psychosis (34)	EAB
Jones, 2010	Psychosis (87)	ESE
Kesting, 2011	Current PDs (28); Remitted PDs (31); Depression (21); Non-clinical (59)	ESE; ISE; SED
Kinderman, 1994	Current PDs (16); Depression (16); Non-clinical (16)	ESE; ISE; SED

<b>Study Ref (First Author, Year)</b>	<b>Participant Group/s<sup>b</sup> (N in Parentheses)</b>	<b>Relevant Domain/s</b>
Kinderman, 1997	Current PDs (20); Depression (20); Non-clinical (20)	EAB
Kinderman, 2003	Current PDs (13); Depression (11); Non-clinical (13)	ESE
Langdon, 2006	Current PDs (19); Non-PD psychosis (15); Non-clinical (21)	EAB
Langdon, 2010	Current PDs (35); Non-clinical (34)	EAB
Langdon, 2013	Current PDs (23); Non-clinical (19)	EAB
Lee, 2004	Current PDs (12); Non-clinical (12)	EAB
Lincoln, 2010	Current PDs (25); Remitted PDs (25); High (25) & low (25) subclinical paranoia	EAB; ESE
Lyon, 1994	Current PDs (14); Depression (14); Non-clinical (14)	EAB; ESE
MacKinnon, 2011	Current PDs (16); Non-clinical (20)	ESE; ISE; SED
Martin, 2002	Current PDs (15); Non-PD psychosis (15); Non-clinical (16)	EAB
McCulloch, 2006	Current PDs (13); Depression (15); Non-clinical (15)	ESE; ISE; SED
McKay, 2005 <sup>c</sup>	Current PDs (13); Remitted PDs (12); Non-clinical (19)	EAB
McKay, 2007 <sup>c</sup>	Current PDs (10); Remitted PDs (10); Non-clinical (19)	ESE; ISE; SED
Mehl, 2010	Current PDs (23); Remitted PDs (18); Non-clinical (22)	EAB
Mehl, 2014 <sup>c</sup>	Psychosis (258); Non-clinical (51)	EAB
Melo, 2006	Current PM PDs (26); Current BM PDs (18); Non-clinical (21)	EAB
Melo, 2013	Current PM PDs (32); Current BM PDs (12); Non-clinical (25)	EAB; ESE
Menon, 2013	Current delusions of reference (18); Non-clinical (17)	EAB
Merrin, 2007	Current PDs (24); Depression (24); Non-clinical (24)	EAB
Mizrahi, 2008	Psychosis (86)	EAB
Moritz, 2006	Current PDs (13); Non-PD psychosis (10); Depression (14); Non-clinical (41)	ESE; ISE; SED
Moritz, 2007	Psychosis (35); Depression (18); Non-clinical (28)	EAB
Palmier-Claus, 2011	Psychosis (256)	SEI
Randall, 2003	Current PDs (18); Remitted PDs (14); Non-clinical (18)	EAB
Randjbar, 2011	Current PDs (10); Non-PD psychosis (19); Non-clinical (33)	ESE
Ringer, 2014	Psychosis (88)	ESE
Romm, 2011	Psychosis (113)	ESE

<b>Study Ref (First Author, Year)</b>	<b>Participant Group/s<sup>b</sup> (N in Parentheses)</b>	<b>Relevant Domain/s</b>
Sharp, 1997	Current delusions (19); Non-PGD psychosis (12); Non-clinical (24)	EAB
Smith, 2005	Current GDs (20); Non-clinical (21)	ESE; ISE; SED
Sundag, 2015 <sup>c</sup>	Current PDs (33); Remitted PDs (10); Non-clinical (33)	ESE
Thewissen, 2008	Current PDs (30); Non-PD Psychosis (34); Remitted psychosis (15); High schizotypy (38); Non-clinical (37)	ESE; SEI
Udachina, 2012	Current PM PDs (14); Current BM PDs (15); Remitted PDs (12); Non-clinical (23)	ESE; SEI
Valiente, 2011	Current PDs (35); Depression (35); Non-clinical (44)	ESE; ISE; SED
Vass, 2015	Psychosis (80)	ESE
Vazquez, 2008	Current PDs (40); Remitted PDs (25); Depression (35); Non-clinical (36)	ESE; ISE; SED
Vorontsova, 2013	Non-depressed PDs (30); Depression (30); Non-clinical (30)	ESE
Warman, 2011	Psychosis (30)	ESE
Wickham, 2015	Psychosis (176)	ESE
Wittorf, 2012	Current PDs (20); Depression (20); Non-clinical (55)	EAB

Abbreviations: BM, bad me; EAB, externalising attributional bias; ESE, explicit self-esteem; GDs, grandiose delusions; ISE, implicit self-esteem; PDs, persecutory delusions; PGDs, persecutory and grandiose delusions; PM, poor me; SED, self-esteem discrepancy; SEI, self-esteem instability.

<sup>a</sup>A more detailed description of the characteristics of these included studies is provided in the supplement.

<sup>b</sup>The participants in the current and remitted delusional groups had psychosis.

<sup>c</sup>Additional data were provided by the authors.

**Table 2. Summary of Meta-Analyses and Meta-Regression Moderator Analyses**

Outcome	N Included Studies	Psychosis, N	Control, N	Hedges' g or r (95% CI)	Heterogeneity: $I^2$ , $\text{Chi}^2$ $P$ -value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, $P$ -value
<b>Externalising attributional bias (EAB)</b>								
Difference in EAB: psychosis with persecutory delusions (PDs) vs non-clinical individuals	27	732	710	$g = 0.48$ (0.23, 0.73)	80%, $P < 0.001$	0.99	Moderate -1 inconsistency	Matching of groups: <sup>a</sup> $N = 16/25$ ; $B = 0.45$ ; $SE = 0.29$ ; $P = 0.113$ Depression differences: <sup>b</sup> $N = 17$ ; $B = 0.05$ ; $SE = 0.22$ ; $P = 0.833$
Difference in EAB: psychosis with PDs vs depression	10	221	200	$g = 1.06$ (0.48, 1.63)	86%, $P < 0.001$	2.15	Moderate -1 inconsistency -1 quality (lack of matching, blinding & power calculations) +1 large effect	—
Difference in EAB: psychosis with PDs vs psychosis without PDs (and, if specified, GDs)	11	232	248	$g = 0.40$ (0.12, 0.68)	53%, $P = 0.018$	-0.38	Moderate -1 imprecision	—
Correlation between EAB and paranoia severity in people with psychosis	21	1128	—	$r = 0.18$ (0.08, 0.27)	58%, $P = 0.001$	0.70	Moderate -1 imprecision	—

Outcome	N Included Studies	Psychosis, N	Control, N	Hedges' g or r (95% CI)	Heterogeneity: $I^2$ , $\chi^2$ $P$ -value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, $P$ -value
<b>Explicit self-esteem (ESE)</b>								
Difference in ESE: psychosis with PDs vs non-clinical individuals	22	576	680	$g = -0.88$ (-1.10, -0.66)	68%, $P < 0.001$	0.18	High	Matching of groups: <sup>a</sup> $N = 12/21$ ; $B = -0.03$ ; $SE = 0.24$ ; $P = 0.910$ Depression differences: <sup>b</sup> $N = 15$ ; $B = -0.70$ ; $SE = 0.23$ ; $P = 0.002$
Difference in ESE: psychosis with PDs vs depression	13	355	292	$g = 0.89$ (0.51, 1.28)	80%, $P < 0.001$	2.05	Moderate -1 inconsistency -1 quality (lack of matching, blinding & power calculations) +1 large effect	Matching of groups: <sup>a</sup> $N = 3/12$ ; $B = -0.49$ ; $SE = 0.50$ ; $P = 0.326$
Difference in ESE: psychosis with PDs vs psychosis without PDs (and, if specified, GDs)	11	411	233	$g = -0.26$ (-0.54, 0.02)	58%, $P = 0.01$	-0.96	Moderate -1 imprecision	—
Correlation between ESE and paranoia severity in people with psychosis	23	1866	—	$r = -0.26$ (-0.34, -0.17)	74%, $P < 0.001$	0.87	High	—

Outcome	N Included Studies	Psychosis, N	Control, N	Hedges' g or r (95% CI)	Heterogeneity: $I^2$ , $\text{Chi}^2$ $P$ -value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, $P$ -value
<b>Implicit self-esteem (ISE)</b>								
Difference in ISE: psychosis with PDs vs non-clinical individuals	11	300	383	$g = -0.37$ (-0.65, -0.08)	66%, $P = 0.001$	-0.06	Low -1 imprecision -1 quality (lack of matching, blinding & power calculations)	Matching of groups: <sup>a</sup> N = 5/11; B = -0.36; SE = 0.28; $P = 0.197$
Difference in ISE: psychosis with PDs vs depression	7	224	174	$g = -0.19$ (-0.45, 0.07)	34%, $P = 0.165$	—	Low -1 imprecision -1 quality (lack of matching, blinding & power calculations)	—
Difference in ISE: psychosis with PDs vs psychosis without PDs (and, if specified, GDs)	4	91	76	$g = -0.24$ (-0.77, 0.30)	61%, $P = 0.054$	—	Low -1 inconsistency -1 imprecision	—
Correlation between ISE and paranoia severity in people with psychosis	4	167	—	$r = -0.13$ (-0.38, 0.15)	62%, $P = 0.049$	—	Low -1 inconsistency -1 imprecision	—

Outcome	N Included Studies	Psychosis, N	Control, N	Hedges' g or r (95% CI)	Heterogeneity: I <sup>2</sup> , Chi <sup>2</sup> P-value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, P-value
<b>Discrepancy scores (DS)<sup>c</sup></b>								
Difference in DS: psychosis with PDs vs non-clinical individuals	10	269	323	g = -0.17 (-0.45, 0.12)	61%, P = 0.006	-0.49	Very low -1 inconsistency -1 imprecision -1 quality (lack of matching, blinding & power calculations)	Matching of groups: <sup>a</sup> N = 5/10; B = 0.07; SE = 0.31; P = 0.823
Difference in DS: psychosis with PDs vs depression	7	224	174	g = 0.61 (0.37, 0.85)	22%, P = 0.258	—	Moderate -1 quality (lack of matching, blinding & power calculations)	—
Difference in DS: psychosis with PDs vs psychosis without PDs (and, if specified, GDs)	4	90	75	g = 0.17 (-0.19, 0.53)	20%, P = 0.287	—	Moderate -1 imprecision	—
Correlation between DS and paranoia severity in people with psychosis	4	165	—	r = 0.09 (-0.09, 0.26)	15%, P = 0.315	—	Moderate -1 imprecision	—
<b>Self-esteem instability (SEI)</b>								
Correlation between SEI and paranoia severity in people with psychosis	4	508	—	r = 0.23 (0.11, 0.34)	38%, P = 0.186	—	High	—

Abbreviations: GDs, grandiose delusions; PDs, persecutory delusions.

<sup>a</sup>'Matching of groups' was a binary moderator (0 = unmatched, 1 = matched). N = number of matched studies/ number of studies that provided information on matching.

<sup>b</sup>'Depression differences' (quantified using the SMD, d) was a continuous moderator. N = number of studies that provided information on depression differences.

Discrepancy scores = scores on discrepancies between implicit and explicit self-esteem.



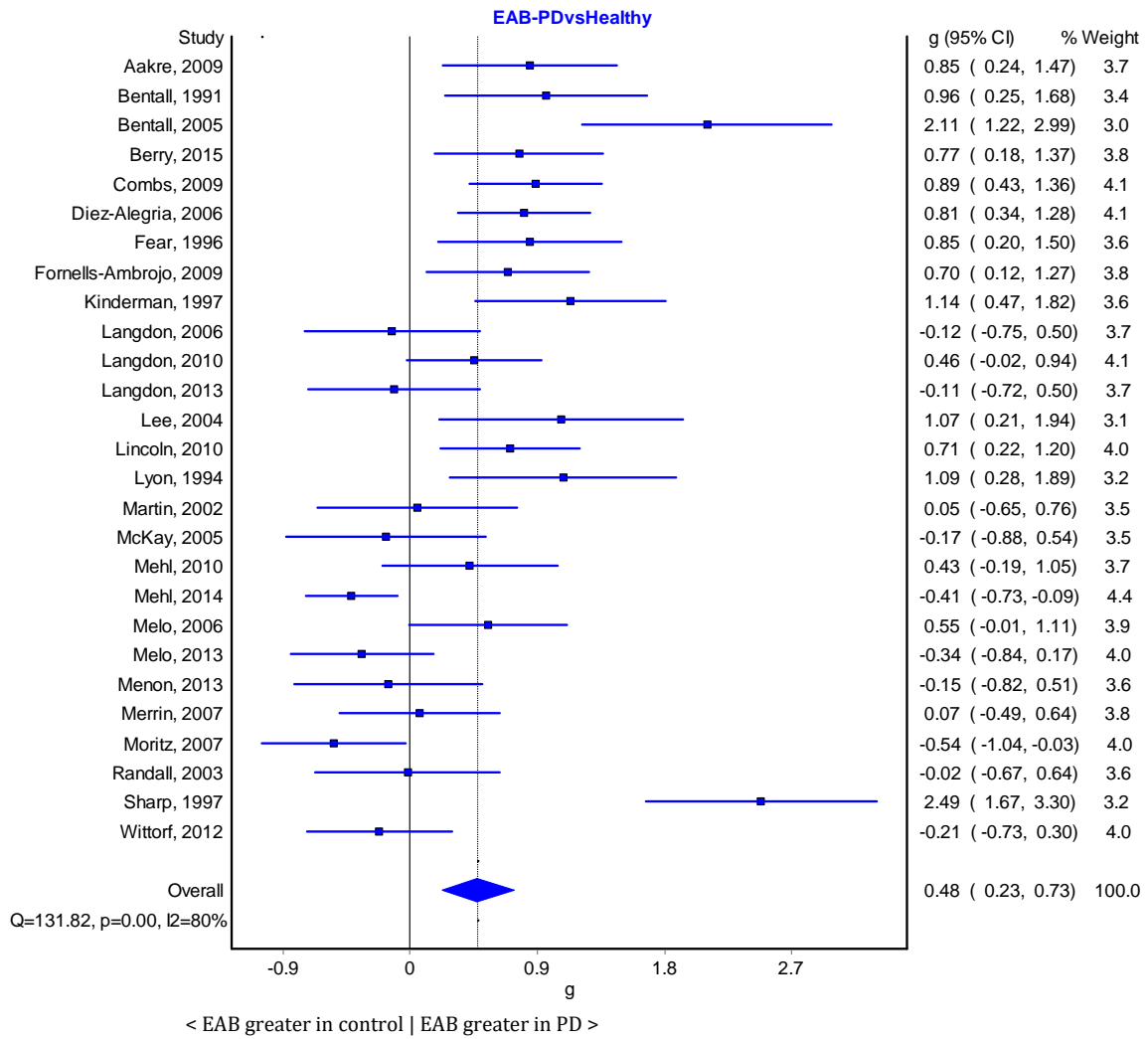
**Table 3. Summary of findings**

<b>Hypothesis</b>	<b>Finding</b>
<b>Externalising attributional bias</b>	
People with persecutory delusions will have a significantly greater externalising attributional bias than non-clinical individuals	Moderate quality evidence in favour of defensive hypothesis (both versions)
People with persecutory delusions will have a significantly greater externalising attributional bias than depressed individuals	Moderate quality evidence in favour of defensive hypothesis (both versions)
People with persecutory delusions will have a significantly greater externalising attributional bias than people with psychosis without persecutory delusions	Moderate quality evidence in favour of defensive hypothesis (both versions)
In people with psychosis, the degree of externalising attributional bias will be positively correlated with paranoia severity	Moderate quality evidence in favour of defensive hypothesis (both versions)
<b>Explicit self-esteem</b>	
People with persecutory delusions will have significantly greater or similar explicit self-esteem compared to non-clinical individuals	High quality evidence against defensive hypothesis (1994 version)
People with persecutory delusions will have significantly greater explicit self-esteem than depressed individuals	Moderate quality evidence in favour of defensive hypothesis (both versions)
People with persecutory delusions will have significantly greater explicit self-esteem than people with psychosis without persecutory delusions	Moderate quality evidence against defensive hypothesis (1994 version)
In people with psychosis, the degree of explicit self-esteem will be positively correlated with paranoia severity	High quality evidence against defensive hypothesis (1994 version)

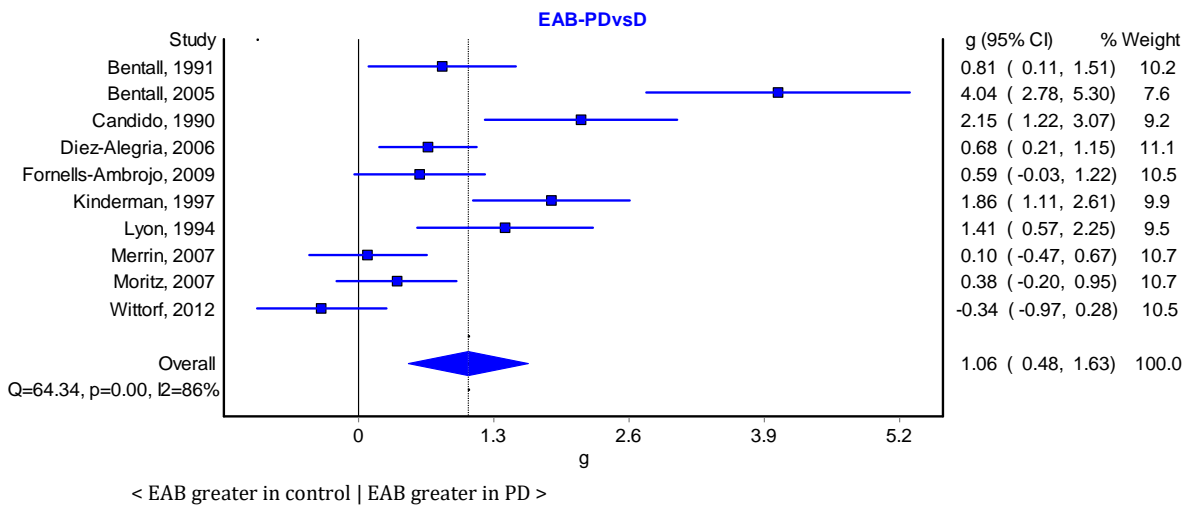
Hypothesis	Finding
<b>Implicit self-esteem</b>	
People with persecutory delusions will have significantly lower implicit self-esteem than non-clinical individuals	Low quality evidence in favour of defensive hypothesis (2001 version)
People with persecutory delusions will have similar implicit self-esteem to depressed individuals	Low quality evidence in favour of defensive hypothesis (2001 version)
People with persecutory delusions will have significantly lower implicit self-esteem than people with psychosis without persecutory delusions	Low quality evidence against defensive hypothesis (2001 version)
In people with psychosis, the degree of implicit self-esteem will be negatively correlated with paranoia severity	Low quality evidence against defensive hypothesis (2001 version)
<b>Self-esteem discrepancy</b>	
People with persecutory delusions will have a significantly greater discrepancy in implicit-explicit self-esteem than non-clinical individuals	Very low quality evidence against defensive hypothesis (2001 version)
People with persecutory delusions will have a significantly greater discrepancy in implicit-explicit self-esteem than depressed individuals	Moderate quality evidence in favour of defensive hypothesis (2001 version)
People with persecutory delusions will have a significantly greater discrepancy in implicit-explicit self-esteem than people with psychosis without persecutory delusions	Moderate quality evidence against defensive hypothesis (2001 version)
In people with psychosis, the degree of implicit-explicit self-esteem discrepancy will be positively correlated with paranoia severity	Moderate quality evidence against defensive hypothesis (2001 version)

<b>Hypothesis</b>	<b>Finding</b>
<b>Self-esteem instability</b>	
In people with psychosis, the degree of self-esteem instability will be positively correlated with paranoia severity.	High quality evidence in favour of defensive hypothesis (2001 version)

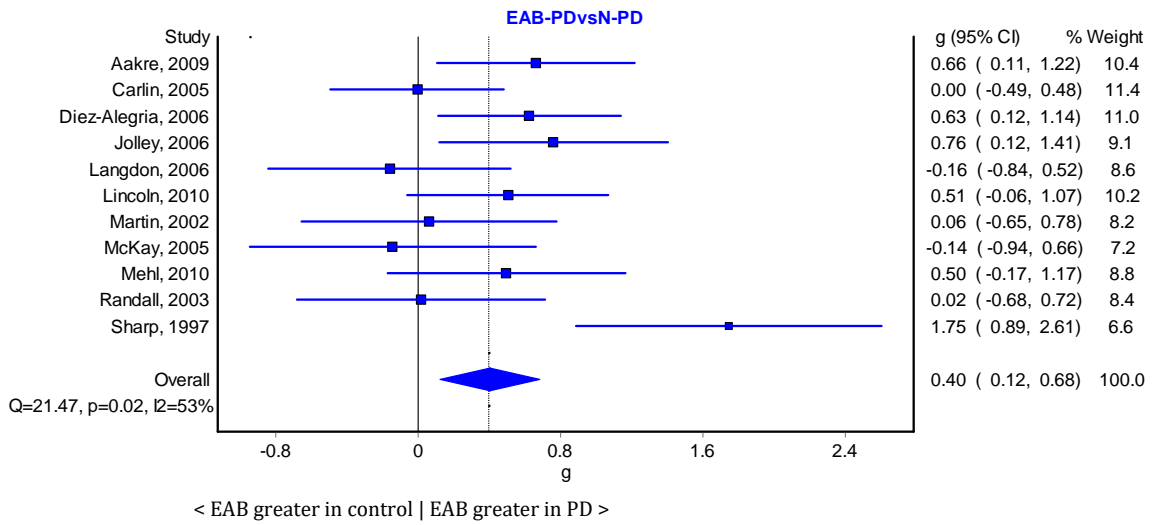
A



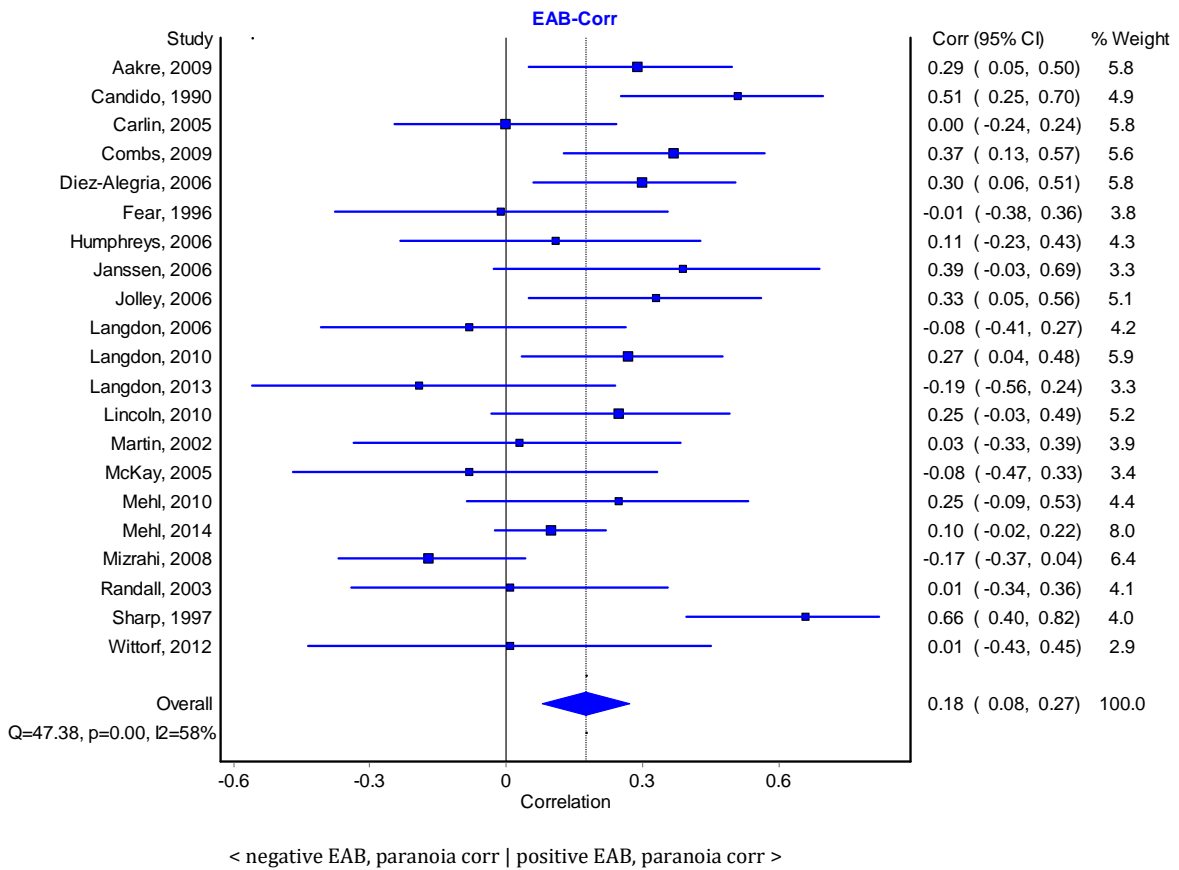
B



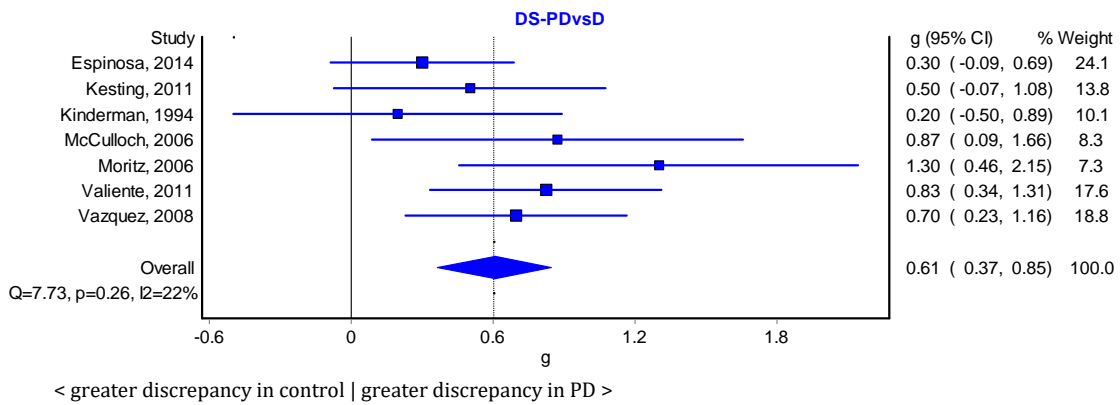
C



D

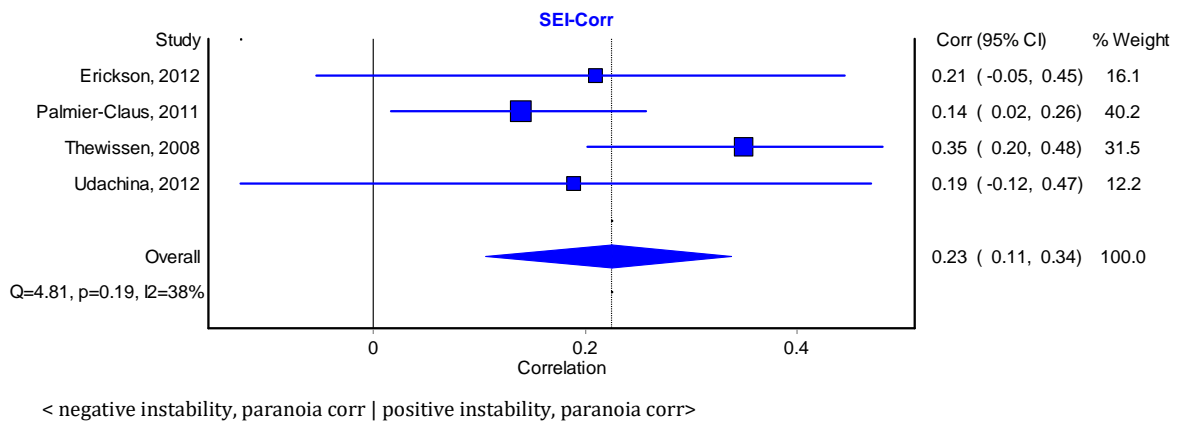


**Figure 5.** Forest plots for analyses of externalising attributional bias (EAB). (A) Forest plot for comparison of EAB between people with psychosis with persecutory delusions (PDs) and healthy (non-clinical) individuals. (B) Forest plot for comparison of EAB between people with psychosis with PDs and people with depression. (C) Forest plot for comparison of EAB between people with psychosis with PDs and people with psychosis without PDs [and, if specified, grandiose delusions (GDs)]. (D) Forest plot of correlation between EAB and paranoia severity in people with psychosis.



**Figure 6.** Forest plot for comparison of discrepancy scores<sup>a</sup> between people with psychosis with persecutory delusions (PDs) and people with depression.

<sup>a</sup>Discrepancy scores = scores on discrepancies between implicit and explicit self-esteem.



**Figure 7.** Forest plot of correlation between self-esteem instability and paranoia severity in people with psychosis.