



This is a repository copy of *The paranoia as defence model of persecutory delusions: a systematic review and meta-analysis*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/141780/>

Version: Accepted Version

Article:

Murphy, P., Bentall, R.P., Freeman, D. et al. (2 more authors) (2018) The paranoia as defence model of persecutory delusions: a systematic review and meta-analysis. *Lancet Psychiatry*, 5 (11). pp. 913-929. ISSN 2215-0366

[https://doi.org/10.1016/S2215-0366\(18\)30339-0](https://doi.org/10.1016/S2215-0366(18)30339-0)

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The paranoia as defence model of persecutory delusions: a systematic review and meta-analysis

Philip Murphy (DClinPsy)^{1,2,3*}
Richard P. Bentall (PhD)⁴
Daniel Freeman (PhD)⁵
Suzanne O'Rourke (PhD)⁶
Paul Hutton (DClinPsy)¹

¹School of Health and Social Care,
Edinburgh Napier University, Edinburgh, UK

²Division of Psychology and Mental Health,
School of Health Sciences,
University of Manchester, Manchester, UK

³Greater Manchester Mental Health NHS Foundation Trust,
Prestwich, Manchester, UK

⁴Department of Psychology,
University of Sheffield, Sheffield, UK

⁵Department of Psychiatry,
University of Oxford, Oxford, UK

⁶Department 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

Citation: Murphy, P., Bentall, R.P., Freeman, D., O'Rourke, S, Hutton, P. The paranoia as defence model of persecutory delusions: a systematic review and meta-analysis. *Lancet Psychiatry*. 2018; published online Oct 9. [http://dx.doi.org/10.1016/S2215-0366\(18\)30339-0](http://dx.doi.org/10.1016/S2215-0366(18)30339-0)

The Lancet Psychiatry homepage: <https://www.journals.elsevier.com/the-lancet-psychiatry>

Summary

Background

An influential psychological model of persecutory delusions proposed that they are caused by a bias towards holding others responsible for negative events (an externalising attributional bias), preventing the individual from becoming aware of underlying low self-esteem. An early version of the model predicted self-esteem would, therefore, be preserved in people with these delusions, but a later version suggested it would be unstable, and that there would be a discrepancy between explicit and implicit self-esteem, with the latter being lower. We did a comprehensive meta-analytical test of the key predictions of this model and assessed the quality of evidence.

Methods

We searched PubMed from Jan 1, 1994, to July 31, 2018, and collated systematic reviews of the defensive model's predictions in relation to persecutory delusions. We also searched PsycINFO, MEDLINE, Embase, and Web of Science for articles published from Jan 1, 2012, to Sept 10, 2016. Cross-sectional data from case-control, longitudinal, or experimental studies that examined self-esteem or the externalising attributional bias in individuals diagnosed as having schizophrenia-spectrum disorder were eligible for meta-analyses of group differences 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 with the Agency for Healthcare Research and Quality assessment tool, and a modified version of Grading of Recommendations Assessment, Development and Evaluation, respectively. The study protocol is registered with PROSPERO, number CRD42016032782.

Findings

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 (21 studies involving 1128 individuals; $r=0.18$, 95% CI 0.08 to 0.27, with moderate quality evidence). People with persecutory delusions also had a greater externalising attributional bias than non-clinical individuals (27 studies involving 1442 individuals; $g=0.48$, 95% CI 0.23 to 0.73) and depressed individuals (ten studies involving 421 individuals; $g=1.06$, 0.48 to 1.63), and people with psychosis without persecutory delusions (11 studies involving 480 individuals; $g=0.40$, 0.12 to 0.68), all based on moderate quality evidence. Contrary to the predictions in the early version of the model, paranoia severity in psychosis was negatively correlated with explicit self-esteem (23 studies involving 1866 individuals; $r=-0.26$, 95% CI -0.34 to -0.17 , with high quality evidence). People with persecutory delusions also had lower explicit self-esteem than non-clinical individuals (22 studies involving 1256 individuals; $g=-0.88$, 95% CI -1.10 to -0.66 , with high quality evidence) and explicit self-esteem similarly low to that in people with psychosis without persecutory delusions (11 studies involving 644 individuals; $g=-0.26$, -0.54 to 0.02, with moderate quality evidence). Consistent with the predictions in the later version of the model, self-esteem instability was positively correlated with paranoia severity in psychosis (four studies involving 508 individuals; $r=0.23$, 95% CI 0.11–0.34, with high quality evidence), and people with persecutory delusions had a greater discrepancy between their implicit and explicit self-esteem than depressed individuals (seven studies involving 398 individuals; $g=0.61$, 95% CI 0.37 to 0.85, with moderate quality evidence). They had higher explicit self-esteem than depressed individuals (13 studies involving 647 individuals; $g=0.89$, 95% CI 0.51 to 1.28, with moderate quality evidence), but similarly low implicit self-esteem (seven studies involving 398 individuals; $g=-0.19$, -0.45 to 0.07, with low quality evidence). In contrast to

the later predictions, people with persecutory delusions did not have a greater self-esteem discrepancy than non-clinical individuals (ten studies involving 592 individuals; $g=-0.17$, 95% CI -0.45 to 0.12), although the evidence was very low quality. People with psychosis with or without persecutory delusions did not differ for implicit self-esteem (four studies involving 167 individuals; $g=-0.24$, 95% CI -0.77 to 0.30 , with low quality evidence) or self-esteem discrepancies (four studies involving 165 individuals; $g=0.17$, -0.19 to 0.53 , with moderate quality evidence).

Interpretation

The predictions that self-esteem would be preserved in people with persecutory delusions in the early version of the paranoia as defence model and that implicit-explicit self-esteem discrepancy would be greater in people with persecutory delusions than in non-clinical individuals and people with psychosis without persecutory delusions in the later version of the model were not supported. By contrast, the later version correctly predicted that people with persecutory delusions have a greater self-esteem discrepancy than people with depression and a greater externalising attributional bias than all control groups, and that both this bias and self-esteem instability are associated with increased paranoia severity. Nevertheless, the reviewed data had limitations. Experimental studies, which might include interventionist-causal trials, are needed.

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, that evaluated the defensive model's predictions in relation to persecutory delusions, with the search term “(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*)”. Only English language studies were considered. We reviewed all papers that referenced either of the two papers introducing each version of the model, and we searched the PROSPERO systematic review database with 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 all argued that there was evidence against both versions. All three, however, acknowledged that many of the individual studies were small and lacked the power to provide precise estimates, or detect theoretically or clinically relevant findings. For the meta-analyses, we collated all the studies cited in the three reviews. Additionally, we searched PsycINFO, MEDLINE, Embase, and Web of Science for papers published from Jan 1, 2012, to Sept 10, 2016, with the search terms used previously. We manually searched the reference lists of all retrieved full-text articles. 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. We found moderate-quality evidence supporting the predictions by both versions of the model that persecutory delusions are associated with an externalising attributional bias. Contrary to the early version predictions, we found evidence of moderate to high quality that people with persecutory delusions have abnormally low explicit self-esteem, and that this is associated with increased severity of paranoia. Supporting the predictions of the later version of the model, we found high-quality evidence that paranoia severity is associated with self-esteem instability and mixed quality evidence that, compared with 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 later version, we found very low quality evidence that such people have a normal, rather than exaggerated, discrepancy in implicit-explicit self-esteem. Comparisons between people with psychosis with and without current persecutory delusions indicate that those with current persecutory delusions have a heightened externalising attributional bias, but group differences in explicit, implicit, and discrepant self-esteem were not evident (evidence was low to moderate quality), 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 but 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 the defensive model, but not others. However, the observational research we reviewed does not allow causal inference. Experimental testing of the model is needed to assess fully the effect of selectively modifying disputed aspects of the model, such as implicit self-esteem, on paranoia and persecutory delusions. For this work, which might at some stage include randomised controlled interventionist-causal trials,

strategies to change these variables in a way that bypasses explicit self-esteem and reliable methods of assessing change in implicit self-esteem will need to be developed.

Introduction

Persecutory (paranoid) delusions involve unfounded beliefs held by individuals that others are trying to harm them¹ and are an important psychiatric problem. Such delusions 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 that they emerge as a consequence of a bias towards holding others responsible for negative events (an externalising attributional bias) to reduce awareness of low self-esteem. In the early (1994) version of the model,⁵ low self-esteem was conceptualised as a discrepancy between one's actual self and one's ideal self (figure 1). Holding others responsible for negative events was proposed to be counterproductive because it might activate fears that others judge the person negatively, in turn increasing the person's attention to interpersonal threat and possibly prompting even more extreme external-personal attributions. Thus, the early version of the model predicted that people with persecutory delusions would have a heightened externalising attributional bias and relatively preserved self-esteem, both of which should be related to increased paranoia severity.

A later version of the model presented in 2001,⁴ known as the Attribution–Self-Representation Cycle model, 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 cyclical process as the individual attempts to explain life events (figure 2). This version suggests that the externalising attributional bias in the context of persecutory delusions provides an incomplete defence against low self-esteem reaching conscious awareness, and that a combination of the externalising attributional bias and low implicit self-esteem will cause inherent instability of self-esteem that will increase with increasing severity of persecutory delusion. Covert measurements of self-esteem and attributions are predicted to 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 of the model are that in people with persecutory delusions, implicit self-esteem, measured by reaction time or similar tasks, will be at a similarly low level to that in people with depression, and there will be a discrepancy between implicit and self-reported explicit self-esteem, with the latter being higher. The later version of the model also considers the origins of external-personal causal inferences, drawing on research which suggests that they involve less cognitive effort than benign external-situational attributions, particularly if an individual has an attentional bias towards threat and difficulty in understanding the intentions of others.

Freeman and colleagues⁶ proposed an alternative non-defensive account of the development and maintenance of persecutory delusions (figure 3), in which persecutory delusions are viewed as threat beliefs, developed in the context of genetic and environmental risk, which are 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 interpersonal experiences, mean that the individual feels inferior to others, different, apart, and, hence, vulnerable. Paranoia feeds on this vulnerability. The model of Freeman and colleagues does not predict a discrepancy between implicit and explicit self-esteem, nor does it claim self-esteem instability is central to persecutory delusion development (although it will be common in people with emotional disorders). The authors regard low self-esteem and negative cognitions as being among several 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 insufficient for persecutory delusions to form, but might be an essential component of one or more complex pathways. These pathways could be sufficient to cause persecutory delusions but are not essential because other complex pathways might also give rise to these delusions.

Given that 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 parsimony⁹ or being difficult to operationalise.⁸ Proponents of the later version of the defensive model place weight on the hypotheses that persecutory delusions involve heightened external-personal attributions, discrepancies between implicit and explicit self-esteem, low implicit self-esteem similar to that in people with depression, and unstable self-esteem, with particular emphasis placed on the latter feature.⁴ Critics, however, argue that even if an external-personal attributional bias is present, its function is moot.⁸ Although the 1994 version of the defensive model predicted that self-esteem would be 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 the later version precludes such predictions, whereas critics have suggested that “relative preservation of mood and (explicit) self-esteem might be expected” even with this defence account.^{7–9}

Three systematic reviews have revealed no clear evidence to support either version of the defensive model and argued that there is evidence against both.^{8,10,11} Each review 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 All the reviews had two notable limitations. First, many studies were small and, therefore, unable to detect reliably the full range of important relationships that might exist,¹² which suggests that meta-analysis is required to form firm conclusions.¹³ Second, all except two assessments of discrepancies between implicit and explicit self-esteem were based on comparisons of results between groups for each type of self-esteem separately.^{14,15} To test the hypothesis of discrepancy adequately, it is argued that the differences between implicit and explicit self-esteem should be compared within groups as well as between groups.^{14,15}

In this study, we aimed to do a systematic review and series of meta-analyses to test key predictions of the early and the later versions of the defensive model. Our approaches were agreed in advance by a leading proponent of the defensive model (RPB) and one of the leading proponents of the non-defensive model (DF). For the early version of the defensive model, we aimed to answer the questions of whether people with persecutory delusions have greater explicit self-esteem than people with depression or with psychosis without persecutory delusions; whether explicit self-esteem is greater than or similar to that of non-clinical individuals; and whether paranoia severity in psychosis is positively correlated with explicit self-esteem. For the later version, we aimed to answer the questions of whether people with persecutory delusions have a greater externalising attributional bias and discrepancy between implicit and explicit self-esteem than 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 to those with depression, but lower implicit self-esteem than nonclinical individuals and people with psychosis without persecutory delusions. We investigated whether correlations would be positive between paranoia severity in psychosis and the degree of externalising attributional bias, implicit-explicit self-esteem discrepancy, and self-esteem instability, and negative between paranoia

severity in psychosis and implicit self-esteem. Finally, we did several prespecified moderator analyses to assess the effects of depression and study quality variables on the overall estimates.

Methods

Search strategy and selection criteria

Three systematic reviews of the relevant literature published in 2013 and 2014 were identified (appendix p 11).^{8,10,11} PsycINFO, MEDLINE, Embase, and Web of Science were then searched by PM (in consultation with PH and a research librarian) for papers published from Jan 1, 2012, to Sept 10, 2016 (appendix p 11). Search terms related to psychosis, delusions, externalising attributional bias, and self-esteem were used. The reference lists of all full-text articles were searched to identify any studies missed in the initial search. Where usable but unpublished data were thought to exist, the relevant authors were contacted. Finally, corresponding authors of all included studies were contacted for any further unpublished data. Studies were selected by PM in consultation with PH. Only English-language studies were considered. Studies were included in the group comparison analyses if they measured externalising attributional bias, implicit self-esteem, or explicit self-esteem in people with a diagnosed schizophrenia spectrum condition (hereafter referred to as psychosis) and those with depression or nonclinical individuals. At least 50% of patients with psychosis had current persecutory delusions, and studies that compared these people with those who had psychosis without persecutory delusions were included in the group comparison analyses unless it was specified that 50% or more of the latter group had current grandiose delusions. Studies without control group data were eligible for inclusion in the correlation analyses if at least 50% of the sample had psychosis and correlation or regression data were reported for paranoia or persecutory ideation and externalising attributional bias or self-esteem. Studies comparing people with current persecutory delusions and 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, except for the self-esteem instability analysis, in which only longitudinal data were used.

We excluded studies where at least 50% of the people with psychosis had bipolar disorder, learning disability, a primary diagnosis of substance-induced psychosis, or psychosis secondary to a general medical condition. When study samples overlapped by 25% or more, we selected the study that reported on the largest number of participants.

As a further step, we searched PubMed for systematic reviews published between Jan 1, 1994, and July 31, 2018, with or without meta-analyses, that evaluated the defensive model's predictions in relation to persecutory delusions. We used the search term “(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 reviewed all papers that referenced either of the two papers introducing each version of the model, and we searched the PROSPERO systematic review database with keywords “paranoia” and “persecutory delusions”. We only found the three previously identified systematic reviews published in 2013 and 2014,^{8,10,11} all of which used narrative synthesis to interpret the evidence, thus confirming the research gap for our current meta-analysis.

The review protocol was reviewed and approved by experts in the psychology of persecutory delusions (DF and RPB) and registered with PROSPERO, number CRD42016032782. All changes to the protocol were decided before analyses were done and are detailed in the appendix (pp 8–10). This study adhered to the statement of PRISMA.¹⁶

Procedures and outcome measures

Different outcomes were selected that corresponded to the different domains of the paranoia as defence model. A data extraction hierarchy was developed for most outcomes that specified which data were preferable and, if these were not reported or could not be acquired, which data would be used. Various scoring methods have been proposed for attributional measures. Therefore, for the externalising attributional bias, the hierarchy was as follows: (1) the external-personal attribution score for negative events (tendency to attribute negative events to other people rather than to oneself or situational factors); (2) the personalising bias score (tendency to attribute negative events to other people rather than to situational factors); (3) the internality attribution score for negative events (tendency to attribute negative events to oneself rather than to other people or situational factors); and (4) the externalising bias score (tendency to attribute negative rather than positive events to external causes, either people or situational factors). Data from the Internal, Personal, and Situational Attributions Questionnaire,¹⁷ which can be used to calculate all four indices in our hierarchy, were preferred over data from the Attributional Style Questionnaire,¹⁸ which can calculate only the latter two indices. Participants' self-ratings were prioritised over ratings given by independent judges.

For explicit self-esteem, we preferred to use data from the Rosenberg Self-Esteem Scale,¹⁹ but when unavailable we used data from conceptually equivalent alternatives. If usable total explicit self-esteem data were unavailable, we prioritised negative explicit self-esteem data over positive explicit self-esteem data. For implicit self-esteem, we used the following hierarchy: (1) Implicit Association Task;²⁰ (2) Emotional Stroop Task;^{21,22} and (3) Go/No-go Association Task.²³ If these data were not available, a conceptually equivalent alternative was used. Implicit and explicit self-esteem discrepancies were calculated from the choice of implicit and explicit self-esteem indices with a method that allowed for the analysis of differences within and between groups, unless already reported (appendix pp 18–21).

Self-esteem instability was assessed by the Experience Sampling Method²⁴ or the repeated application of a self-esteem measure, such as the Rosenberg Self-Esteem Scale. Full details about the selection of outcome measures are given in the appendix (pp 15–17).

Data were extracted into a spreadsheet by PM and crosschecked by PH. Means and SDs were used for analyses of group differences. Missing SDs were, where possible, calculated from *t* test values, *p* values, *F* values, SEs or CIs, with equations in the Cochrane Handbook²⁵ and elsewhere.²⁶ Alternatively, we estimated SDs from the mean SD of the other included studies.²⁷ For within-group analyses, correlation coefficients were extracted directly from the results of a study. If correlation coefficients were not reported, they were derived in one of two ways: from group differences between people with psychosis with and without current persecutory delusions with the Campbell Collaboration effect size calculator,²⁸ or from regression coefficients.^{29,30}

Statistical analysis

Meta-analyses were done with MetaXL software (version 5.3). For group difference meta-analyses, Hedges' *g* standardised mean differences and 95% CIs were computed. When a study had two or more similar groups, these were combined with equations specified in the Cochrane Handbook.²⁵ For correlational meta-analyses, Pearson's correlations were converted into Fisher's *Z* scores and 95% CIs, as were any Spearman correlations after first being converted into Pearson's correlations.³¹ The meta-analytical estimates were then back-transformed into Pearson's correlations to allow interpretation. Following Cohen's conventions,³² Hedges' *g* values of 0.2, 0.5, and 0.8 were interpreted as small, moderate, and large group differences,

respectively, and Pearson's r values of 0.1, 0.3, and 0.5 were interpreted as small, moderate, and large correlations, respectively.

Random-effects meta-analyses were done with the DerSimonian and Laird method³³ for all outcomes.³² When heterogeneity was less than moderate (ie, $I^2 < 40\%$),²⁵ we assessed sensitivity with a fixed-effect analysis,³⁴ but as we found no substantively different findings, we do not report these data. Publication bias was assessed through the Doi plot and Luis Furuya-Kanamori asymmetry (LFK) index for outcomes based on at least ten studies,²⁵ because this method is more sensitive than the funnel-plot method.³⁵ We also created funnel plots and used Egger's test, but the latter was reported only where it differed from the LFK value. The "trim and fill" method was applied if the LFK index indicated bias (LFK > 2).³⁶

We assessed two prespecified moderators of effect size: the matching of groups by demographics (age, sex, education [or a measure of intelligence quotient if education was not reported], and ethnicity), and group differences in depression (appendix p 22). If ten or more studies in a meta-analysis provided usable data, random effects meta-regression was used to test these moderator effects with Comprehensive Meta-Analysis software (version 3).²⁵ Two moderator analyses, the blinding of the outcome assessor and early versus chronic psychosis, were abandoned because of insufficient data (fewer than five studies per level of variable).

In one meta-regression moderator analysis, group differences in depression significantly moderated an effect size. To explore further the effect of depression on the relevant effect size, we conducted a subgroup analysis. We coded groups of people with persecutory delusions as either depressed (\geq mild depression) or non-depressed ($<$ mild depression) based on a cutoff score on a reported measure of depression (appendix p 22), and a mixed effects analysis using Comprehensive Meta-Analysis software (version 3) was done.

The methodological quality of all studies was assessed with an adapted version of the Agency for Healthcare Research and Quality assessment tool (appendix pp 61–63).³⁷ The quality of the meta-analytical outcomes was assessed with an adapted version of the Grading of Recommendations Assessment, Development and Evaluation approach (appendix p 95).³⁸ The overall rating in this system (high, moderate, low, or very low quality) incorporates quality of the studies, publication bias, inconsistency, and imprecision.

Role of the funding source

There was no funding source for this study. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Of 104 studies with full-text reports that were assessed for eligibility, 40 were excluded (figure 1, appendix pp 13–14). 64 studies were included in the analysis, among which 33 tested hypotheses on the externalising attributional bias, 36 on explicit self-esteem, 11 on implicit self-esteem, ten on self-esteem discrepancies, and four on self-esteem instability (figure 4, table 1, appendix pp 23–41). Unpublished data were obtained from the authors of six studies.^{39–44} The included studies were published between 1991 and 2016, 32 had been done in the UK and the remainder had been done in Europe (17), the USA and Canada (ten), and Australia (five).

Consistent methodological problems were nonreporting of prespecified power calculations and nonblinding of researchers to diagnosis. Groups were not well matched for important demographic variables. Selection of participants was generally unbiased, although convenience samples were widely employed. The studies generally provided adequate sample characteristics and used valid and reliable measures of diagnostic status, persecutory delusion severity, and self-esteem, but just over a third of the externalising attributional bias measures were judged to be only partly reliable and valid, primarily because they represented the bottom two data extraction hierarchy indices (ie, they did not distinguish between external-personal and external-situational attributions; table 2, appendix pp 65–93).

As predicted by both versions of the paranoia as defence model, people with persecutory delusions had a significantly greater externalising attributional bias than nonclinical individuals and people with psychosis without persecutory delusions (table 2, figure 5). Externalising differed substantially between people with persecutory delusions and those with depression, with the former having an exaggerated bias (table 2, figure 5). There was a small significant positive correlation between paranoia severity and the externalising attributional bias in psychosis (table 2, figure 5).

Consistent with the 1994 version of the paranoia as defence model, people with persecutory delusions had significantly greater explicit self-esteem than people with depression but, contrary to its predictions, people with persecutory delusions had significantly lower explicit self-esteem than non-clinical individuals and similar explicit self-esteem to people with psychosis without persecutory delusions (table 2). Also contrary to the 1994 version, we calculated a small to moderate significant negative correlation between paranoia severity and explicit self-esteem in psychosis (table 2).

Consistent with the predictions of the 2001 version of the model, people with persecutory delusions had significantly lower implicit self-esteem than non-clinical individuals and similar implicit self-esteem to people with depression (table 2). However, inconsistent with this version was no significant difference in implicit self-esteem between people with psychosis with and without persecutory delusions and no significant correlation between paranoia severity and implicit self-esteem in psychosis (table 2).

As predicted by the 2001 version of the paranoia as defence model, people with persecutory delusions had a significantly greater discrepancy between their implicit and explicit self-esteem than people with depression (table 2, figure 6), but there was no evidence that people with persecutory delusions had a significantly greater implicit-explicit self-esteem discrepancy than non-clinical individuals or people with psychosis without persecutory delusions (table 2). No significant correlation was found between paranoia severity and discrepancy scores in psychosis (table 2).

As predicted by the 2001 version of the model, there was a significant and positive correlation between paranoia severity and self-esteem instability in psychosis (table 2, figure 7). Group differences in self-esteem instability were unavailable.

Differences in severity of depression in people with psychosis 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 ($Q=31.71$, $p=0.003$), which suggests that there is unexplained variance in explicit self-esteem group differences. No other moderator analyses were significant (table 2).

A subgroup analysis on the explicit self-esteem data showed that individuals with persecutory delusions and depression had lower explicit self-esteem than non-clinical individuals (12 studies involving 698 individuals; $g=-0.99$, 95% CI -1.28 to -0.70 , $Z=-6.71$, $p<0.001$), but that explicit self-esteem did not differ between those with persecutory delusions who were not depressed and non-clinical individuals (five studies involving 296 individuals; $g=-0.51$, -1.09 to 0.08 , $Z=-1.69$, $p=0.091$). However, the difference between the two effect sizes was not significantly different ($Q=2.09$, $p=0.148$).

Potential publication bias was indicated for the analyses of externalising attributional bias and explicit self-esteem (psychosis with persecutory delusions vs depression; table 2, appendix pp 106–110). However, the “trim and fill” method³⁴ did not impute any missing studies and, therefore, the point estimates remained the same.

The quality of evidence at the outcome level was generally moderate to high (tables 2, 3). All the evidence for implicit self-esteem outcomes was of low quality, and quality of evidence was very low for one of the implicit-explicit self-esteem discrepancy outcomes (psychosis with persecutory delusions vs non-clinical individuals; tables 2, 3).

Discussion

Over the past 25 years, 64 studies involving 5363 participants (3562 participants with psychosis, 442 participants with depression, and 1359 non-clinical participants) have tested the paranoia as defence model of persecutory delusions. By doing a meta-analytical appraisal of published and unpublished evidence for the 1994 and 2001 versions of the model, we were able to overcome the power limitations of individual studies. We also calculated the discrepancy between implicit and explicit self-esteem to enable analysis of differences within and between groups. Our study protocol was approved by exponents of the general paranoia as defence model^{4,5} and of 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 increased externalising attributional bias. This bias seems to be specific to persecutory delusions and is associated with paranoia severity. The predictions of the 2001 version of the model were supported by an observed association between self-esteem instability and paranoia severity and our finding that individuals 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 a relative implicit-explicit self-esteem discrepancy, which seemed to be confirmed in a direct comparison of self-esteem discrepancies. Although explicit self-esteem was considerably lower among people with persecutory delusions than among non-clinical individuals, the moderator analysis suggests that this difference could be at least partly a function of co-occurring depression.

Critics of the paranoia as defence model might reasonably note that an increased externalising attributional bias does not in itself indicate anything about its function.^{6,8,9} The conceptual overlap between holding others responsible for negative events and worrying that others wish to cause one harm might also be concerning.^{6,45} In relation to the 2001 version of the model, although critics might acknowledge that self-esteem instability is linked to paranoia severity, they might see no need to invoke defensive explanations and query the specificity of its effects to paranoia. They might suggest that the low level of implicit self-esteem 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. Critics might query whether the pattern of self-esteem findings is attributable to the characteristics of people with persecutory delusions, or whether they reveal more about the self-esteem profile of people with depression. Moreover, the specific claim of an exaggerated self-esteem discrepancy in people with persecutory delusions in the 2001 version of the paranoia as defence model was not supported by the evidence. Thus, critics of the defensive account might argue that a non-defensive account reflects a more parsimonious interpretation of the pattern of findings when comparisons with non-clinical individuals are considered. It could also be said that although explicit self-esteem might be higher when people with persecutory delusions are not depressed, this pattern applies to relatively few people with such difficulties,^{46,47} which has been claimed to be inconsistent with a defensive model, or at the least the 1994 version. Implicit or discrepant self-esteem did not differ between individuals with psychosis with persecutory delusions and psychosis without persecutory delusions, which casts 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 defensive model.

In response to the concern about the conceptual overlap between the externalising attributional bias and paranoia severity, proponents of the defensive model could argue that the correlation between the externalising attributional bias and paranoia severity that we found was only small

in magnitude, which is inconsistent with a substantial conceptual overlap. Proponents of the defensive model could argue that the 2001 version successfully predicted that self-esteem instability would be associated with paranoia severity⁴ and, although there could indeed be non-defensive explanations for this relationship, a strength of the model is its ability to make specific predictions that are subsequently supported by evidence. Although the relationship between explicit self-esteem and paranoia does not require invocation of defensive processes, neither does it preclude them. Indeed, the non-defensive model and the 2001 version of the defensive model both predict a reciprocal relationship between paranoia and low explicit self-esteem. In relation to the self-esteem discrepancy findings, how the nature of the discrepancy in non-clinical and depressed individuals is viewed is crucial to interpretation. If non-clinical individuals have a self-esteem discrepancy and depressed individuals do not, it follows from our meta-analytical findings 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 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 fulfils its defensive function (ie, it does not fully preserve explicit self-esteem but prevents explicit self-esteem from falling to the even lower level of implicit self-esteem).⁴⁸

Although many people with persecutory delusions are indeed depressed,^{46,47} there is also evidence that fluctuations in mood are strongly associated with the formation and maintenance of paranoia in the general population,⁴⁹ which proponents could point out would be consistent with the predictions in the 2001 version of the paranoia as defence model. Moreover, if depression is common in people with persecutory delusions, then an adequate test of the defensive model would need to control for this in some way, since depression might indicate that defensive processes do not adequately maintain a self-esteem discrepancy. Thus, if people with persecutory delusions are generally depressed, then comparisons with people with depression alone might be more informative about what is specific to persecutory delusions. Indeed, because depression is increased in people with persecutory delusions, it is unclear why they do not have a much smaller discrepancy between implicit and explicit self-esteem when compared with non-clinical individuals.

Although our findings could 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 people with psychosis without current persecutory delusions could plausibly continue to carry the self-esteem risk factors that make them vulnerable to developing persecutory delusions. Proponents of the defensive model might suggest that the 2001 version represents an “unnecessary but sufficient” cause of persecutory delusions,⁵⁰ in which case 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” components of this process,⁵⁰ and, therefore, might all be required for persecutory delusion occurrence, something which few studies have measured. Finally, heterogeneity in meta-analytical estimates is often taken to reflect the presence of unknown moderators and, thus, might be viewed as informative. For instance, variance in cross-sectional estimates of self-esteem in persecutory delusions first motivated researchers to investigate whether self-esteem instability might be tied to paranoia severity.⁴

A further point by critics of the defensive model could be that the small 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. Additionally, critics might 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 if the defensive account cannot detect self-esteem differences between people with current persecutory delusions and those with psychosis without persecutory delusions (even those who have had such difficulties in the past), another model would be needed to account for delusion occurrence. Fluctuations in self-esteem, emotional dysregulation, and mood instability are common in many mental health disorders, especially where negative self-views are involved (eg, depression, eating disorders, and borderline personality disorders), and these, it could be argued, are part of emotional difficulties and not a sign of defence processes. Negative experiences are likely to trigger fluctuations in emotional state, and mood instability can occur in hallucinations⁵¹ and is associated with susceptibility to depression.⁵² A strength of a model, critics might conclude, is not just its predictive power, but also its parsimony – ie, that all findings can be explained without the need to invoke defensive processing.

Our meta-analysis clearly leaves much room for continued disagreement. Our findings do, however, show an association between persecutory delusions and an externalising attributional bias. 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 paranoia as defence model will note that the interpretation of their overall pattern depends on whether non-clinical individuals do have a discrepancy and whether people with depression do not. A complicating factor in the resolution of this debate is the concern over the validity of measures of implicit self-esteem.⁵³ Noting this, Buhrmester and colleagues⁵³ have argued for a different approach to measuring this construct: “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.”

We also found that people with persecutory delusions have abnormally low self-esteem (whether explicit or implicit) compared with non-clinical individuals. Thus, the hypothesis of the 1994 version of the paranoia as defence model, that self-esteem is preserved by the operation of defensive attributional processes, can be rejected. Moreover, mounting experimental evidence indicates that selectively improving explicit self-esteem can cause improvements in clinical and non-clinical paranoia.^{54–56} Thus, regardless of whether defensive processes are also at play, negative explicit self-esteem is likely to contribute directly to paranoia.

Observational research generally precludes firm causal inferences, even with the increased power afforded by meta-analysis. Thus, the 2001 version of the paranoia as defence model needs to undergo experimental testing 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 might include randomised controlled interventionist-causal trials,^{57,58} 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 substantial and selective improvement in implicit self-esteem can be achieved, and if this causes an improvement in persecutory delusions and self-esteem stability, it would be strong support for the paranoia as defence model.

Resource constraints meant that we were limited to English-language studies, but it is unlikely that data excluded for this reason would have had a substantial effect upon effect sizes or our conclusions. The number of studies was insufficient for some of the planned moderator analyses and tests of publication bias. Additionally, although most of the analyses produced reasonable quality evidence, the estimates for implicit self-esteem had low reliability, partly because of the methodological difficulty in measuring this characteristic.^{10,11,59} Finally, the complexity of this review made it challenging to minimise the time-lag between search completion and publication. We are, however, unaware of any major new studies having been published since the initial search was completed.

The paranoia as defence model is an influential model of persecutory delusions^{4,5} which proposes that 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 that supports several predictions of the 2001 version of the model, but also some evidence that does not. Our findings indicate that the prediction in the 1994 version that persecutory delusions will involve preserved self-esteem can be rejected. Experimental research that manipulates the key variables of interest should now be done to resolve the debate and determine whether the defensive model has clinical implications for people with persecutory delusions.

Contributors

PM and PH conceived the study. PM wrote the first draft of the protocol, RPB, DF, and PH substantially contributed to its critical revision, and PM, RPB, DF, and PH read and approved the final protocol. PM did 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 SO'R substantially contributed to the interpretation of the data and the critical revision of the manuscript. All authors read and approved the final manuscript.

Declaration of interests

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 a cofounder of Oxford VR, a University of Oxford spinout company, which builds psychological therapies with immersive technology. The other authors declare no competing interests.

Acknowledgments

We thank the peer reviewers who provided valuable input into the drafting of the final manuscript. We also thank the authors who provided unpublished data.

References

- 1 Freeman D, Garety PA. Comments on the content of persecutory delusions: does the definition need clarification? *Br J Clin Psychol* 2000; 39: 407–14.
- 2 Coid JW, Ullrich S, Kallis C, et al. The relationship between delusions and violence. *JAMA Psychiatry* 2013; 70: 465–71.
- 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* 1994; 165: 101–06.
- 4 Bentall RP, Corcoran R, Howard R, Blackwood N, Kinderman P. Persecutory delusions: a review and theoretical integration. *Clin Psychol Rev* 2001; 21: 1143–92.
- 5 Bentall RP, Kinderman P, Kaney S. The self, attributional processes and abnormal beliefs: towards a model of persecutory delusions. *Behav Res Ther* 1994; 32: 331–41.
- 6 Freeman D, Garety PA, Kuipers E, Fowler D, Bebbington PE. A cognitive model of persecutory delusions. *Br J Clin Psychol* 2002; 41: 331–47.
- 7 Freeman D. Persecutory delusions: a cognitive perspective on understanding and treatment. *Lancet Psychiatry* 2016; 3: 685–92.
- 8 Garety PA, Freeman D. The past and future of delusions research: from the inexplicable to the treatable. *Br J Psychiatry* 2013; 203: 327–33.
- 9 Freeman D, Garety P. Advances in understanding and treating persecutory delusions: a review. *Soc Psychiatry Psychiatr Epidemiol* 2014; 49: 1179–89.
- 10 Kesting M-L, Lincoln TM. The relevance of self-esteem and self-schemas to persecutory delusions: a systematic review. *Compr Psychiatry* 2013; 54: 766–89.
- 11 Tiernan B, Tracey R, Shannon C. Paranoia and self-concepts in psychosis: a systematic review of the literature. *Psychiatry Res* 2014; 216: 303–13.
- 12 Maxwell SE. The persistence of underpowered studies in psychological research: causes, consequences, and remedies. *Psychol Methods* 2004; 9: 147–63.
- 13 Borenstein M, Hedges L, Higgins J, Rothstein H. Vote counting— a new name for an old problem. In: introduction to meta-analysis. Chichester: Wiley, 2009: 251–55.
- 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* 2011; 186: 197–202.
- 15 Vázquez C, Díez-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* 2008; 39: 587–99.

- 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* 2009; 151: 264–69.
- 17 Kinderman P, Bentall R. A new measure of causal locus: the Internal, Personal and Situational Attributions Questionnaire. *Pers Individ Dif* 1996; 20: 261–64.
- 18 Peterson C, Semmel A, von Baeyer C, Abramson LY, Metalsky GI, Seligman MEP. The Attributional Style Questionnaire. *Cognit Ther Res* 1982; 6: 287–99.
- 19 Rosenberg M. Rosenberg self-esteem scale (RSE). In: Ciarrochi J, Bilich L. *Acceptance and commitment therapy*. Wollongong, NSW: University of Wollongong, 1965: 61–62.
- 20 Greenwald AG, McGhee DE, Schwartz JLK. Measuring individual differences in implicit cognition: the implicit association test. *J Pers Soc Psychol* 1998; 74: 1464–80.
- 21 Williams JMG, Mathews A, MacLeod C. The Emotional Stroop Task and psychopathology. *Psychol Bull* 1996; 120: 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* 2001; 19: 625–66.
- 24 Csikszentmihalyi M, Larson R. Validity and reliability of the experience-sampling method. *J Nerv Ment Disord* 1987; 175: 526–36.
- 25 Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. March, 2011. <http://handbook-5-1.cochrane.org> (accessed Sept 10, 2018).
- 26 Borenstein M, Hedges L, Higgins J, Rothstein H. Converting among effect sizes. In: *Introduction to meta-analysis*. Chichester: Wiley, 2009: 45–49.
- 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* 2006; 59: 7–10.
- 28 Wilson D. Practical meta-analysis effect size calculator. 2017. <https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php> (accessed Sept 10, 2018).
- 29 Peterson RA, Brown SP. On the use of beta coefficients in meta-analysis. *J Appl Psychol* 2005; 90: 175–81.
- 30 Kelley K, Maxwell SE. Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant. *Psychol Methods* 2003; 8: 305–21.
- 31 Rupinski M, Dunlap W. Approximating Pearson product-moment correlations from Kendall's tau and Spearman's rho. *Educ Psychol Meas* 1996; 56: 419–29.
- 32 Cohen J. *Statistical power analysis for the behavioural sciences*, 2nd edn. New York, NY: Academic Press. 1988: 567.

- 33 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7: 177–88.
- 34 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; 1: 97–111.
- 35 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* 2018; published online April 4. DOI:10.1097/ XEB.000000000000141.
- 36 Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; 56: 455–63.
- 37 Williams J, Plassman B, Burke J, Benjamin S, Benjamin S. Evidence report/technology assessment number 193. Preventing Alzheimer’s disease and cognitive decline. April, 2010. <https://www.ahrq.gov/downloads/pub/evidence/pdf/alzheimers/alzcog.pdf> (accessed Sept 10, 2018).
- 38 Guyatt G, Oxman A, Vist G, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Br Med J* 2008; 336: 1106–10.
- 39 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* 2015; 226: 84–90.
- 40 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* 2009; 48: 141–62.
- 41 McKay R, Langdon R, Coltheart M. Paranoia, persecutory delusions and attributional biases. *Psychiatry Res* 2005; 136: 233–45.
- 42 McKay R, Langdon R, Coltheart M. The defensive function of persecutory delusions: an investigation using the Implicit Association Test. *Cogn Neuropsychiatry* 2007; 12: 1–24.
- 43 Mehl S, Landsberg MW, Schmidt A-C, et al. Why do bad things happen to me? Attributional style, depressed mood, and persecutory delusions in patients with schizophrenia. *Schizophr Bull* 2014; 40: 1338–46.
- 44 Sundag J, Lincoln TM, Hartmann MM, Moritz S. Is the content of persecutory delusions relevant to self-esteem? *Psychosis* 2015; 7: 237–48.
- 45 Freeman D, Garety P. Advances in understanding and treating persecutory delusions: a review. *Soc Psychiatry Psychiatr Epidemiol* 2014; 49: 1179–89.
- 46 Buckley PF, Miller BJ, Lehrer DS, Castle DJ. Psychiatric comorbidities and schizophrenia. *Schizophr Bull* 2009; 35: 383–402.

- 47 Vorontsova N, Garety P, Freeman D. Cognitive factors maintaining persecutory delusions in psychosis: the contribution of depression. *J Abnorm Psychol* 2013; 122: 1121–31.
- 48 Garety PA, Freeman D. Cognitive approaches to delusions: a critical review of theories and evidence. *Br J Clin Psychol* 1999; 38: 113–54.
- 49 Marwaha S, Broome MR, Bebbington PE, Kuipers E, Freeman D. Mood instability and psychosis: analyses of British national survey data. *Schizophr Bull* 2014; 40: 269–77.
- 50 Mackie JL. *The cement of the universe: a study of causation*. Oxford: Oxford University Press, 1974.
- 51 Marwaha S, Bebbington P. Mood as a mediator of the link between child sexual abuse and psychosis. *Soc Psychiatry Psychiatr Epidemiol* 2015; 50: 661–63.
- 52 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; 45: 1531–41.
- 53 Buhrmester MD, Blanton H, Swann WB. Implicit self-esteem: nature, measurement, and a new way forward. *J Pers Soc Psychol* 2011; 100: 365–85.
- 54 Freeman D, Pugh K, Dunn G, 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: 186–92.
- 55 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; 218: 348–52.
- 56 Atherton S, Antley A, Evans N, et al. Self-confidence and paranoia: an experimental study using an immersive virtual reality social situation. *Behav Cogn Psychother* 2016; 44: 56–64.
- 57 Kendler KS, Campbell J. Interventionist causal models in psychiatry: repositioning the mind–body problem. *Psychol Med* 2009; 39: 881–87.
- 58 Freeman D, Dunn G, Startup 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. *Lancet Psychiatry* 2015; 2: 305–13.
- 59 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* 2000; 79: 631–43.

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

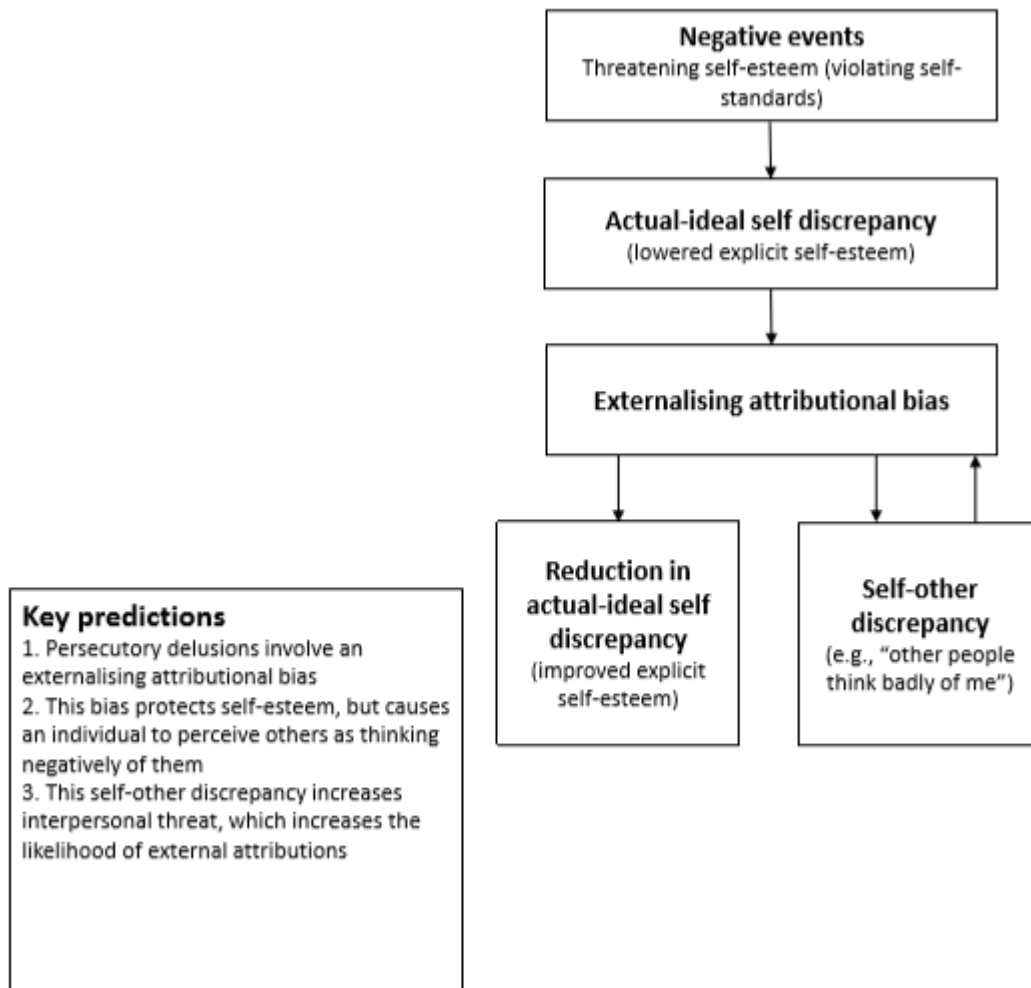


Figure 2. The revised ‘paranoia-as-defence’ model: Bentall et al. (2001)

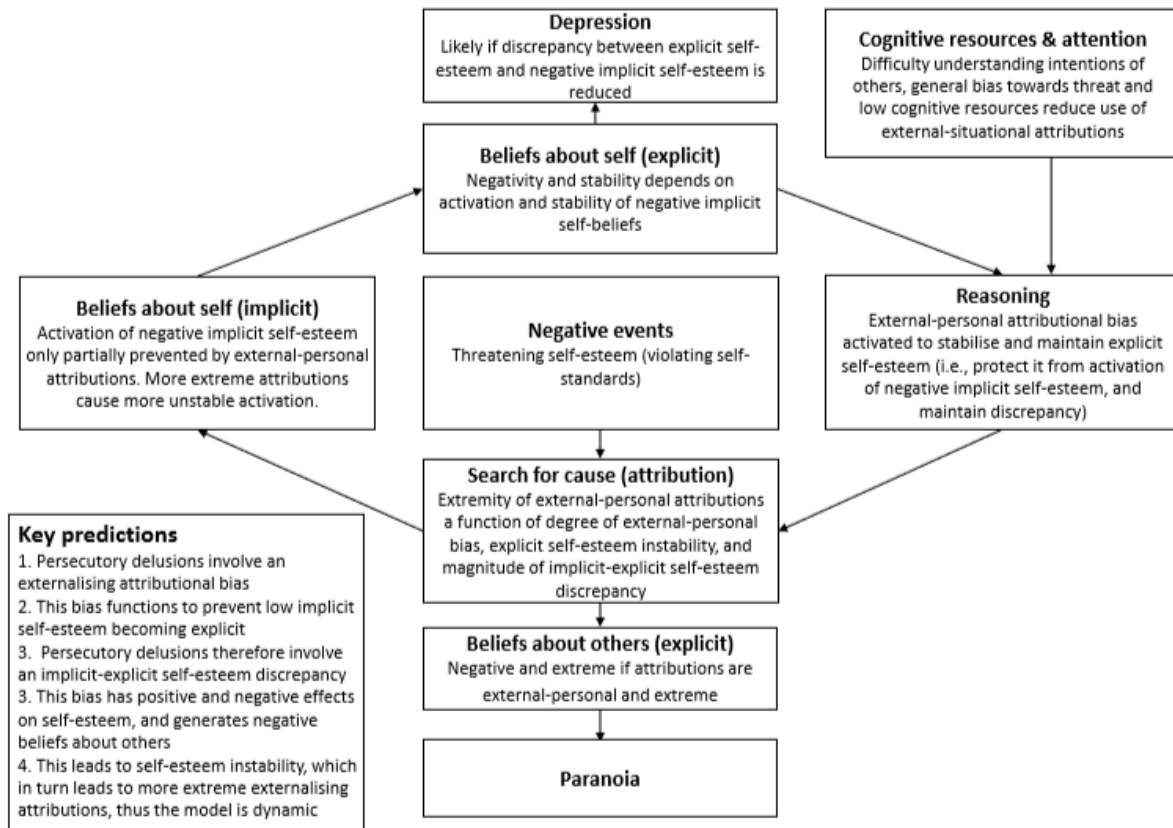
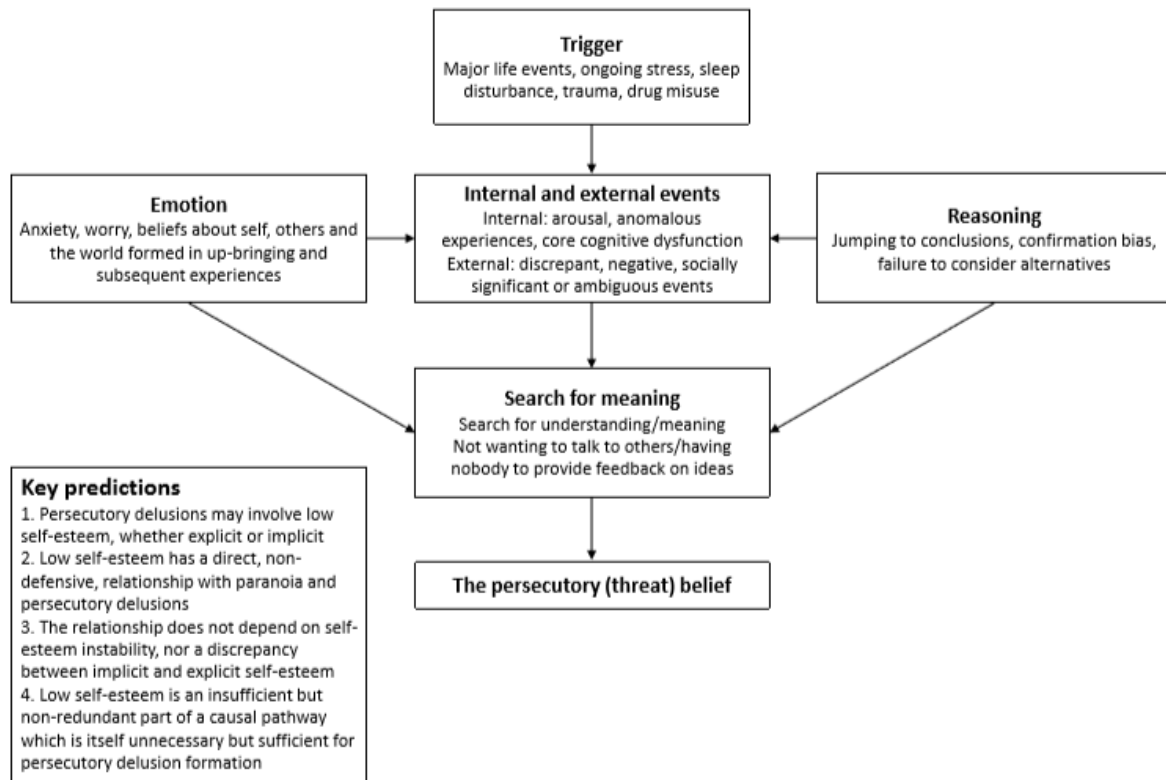


Figure 3. A non-defensive cognitive model of persecutory delusions: Freeman et al. (2002)



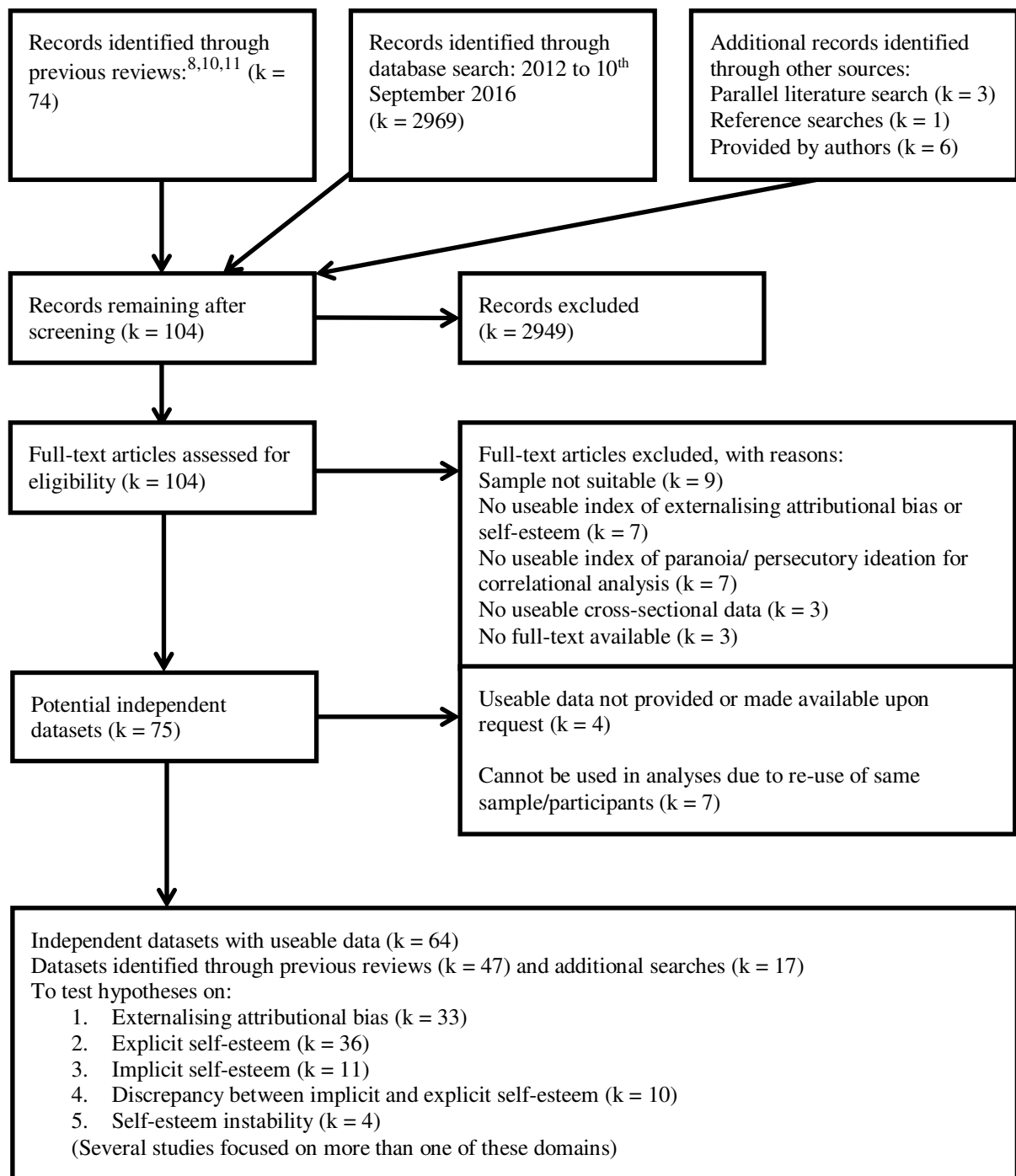


Figure 4. PRISMA flowchart of study selection

Table 1. Studies included in the meta-analysis^a

Study Ref (First Author, Year)	Participant Group/s^b (N in Parentheses)	Relevant Domain/s
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 ^c	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 ^c	Current PM PDs (20); Depression (21); Non-clinical (32)	EAB; ESE
Freeman, 1998	Current PDs (28); Non-PD delusions (25)	ESE
Freeman, 2013	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

Study Ref (First Author, Year)	Participant Group/s^b (N in Parentheses)	Relevant Domain/s
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 ^c	Current PDs (13); Remitted PDs (12); Non-clinical (19)	EAB
McKay, 2007 ^c	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 ^c	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

Study Ref (First Author, Year)	Participant Group/s^b (N in Parentheses)	Relevant Domain/s
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 ^c	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.

^aMore details and references of the studies included in the meta-analysis are provided in the appendix (pp 23-41).

^bThe participants in the current and remitted delusional groups had psychosis.

^cAdditional 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: I², Chi² P-value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, P-value
Externalising attributional bias (EAB)								
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: ^a N = 16/25; B = 0.45; SE = 0.29; P = 0.113 Depression differences: ^b 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 , χ^2 P -value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, P -value
Explicit self-esteem (ESE)								
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: ^a N = 12/21; $B = -0.03$; SE = 0.24; $P = 0.910$ Depression differences: ^b 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: ^a 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 , χ^2 P -value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, P -value
Implicit self-esteem (ISE)								
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: ^a 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^2 , χ^2 P -value	Publication bias: LFK index	Quality (GRADE)	Moderator: N, B, SE, P -value
Discrepancy scores (DS)^c								
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: ^a $N = 5/10$; $B = 0.07$; $SE = 0.31$; $P = 0.823$
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	—
Self-esteem instability (SEI)								
Correlation between SEI and paranoia severity in people with psychosis	4	508	—	$r = 0.23$ (0.11, 0.34)	38%, $P = 0.186$	—	High	—

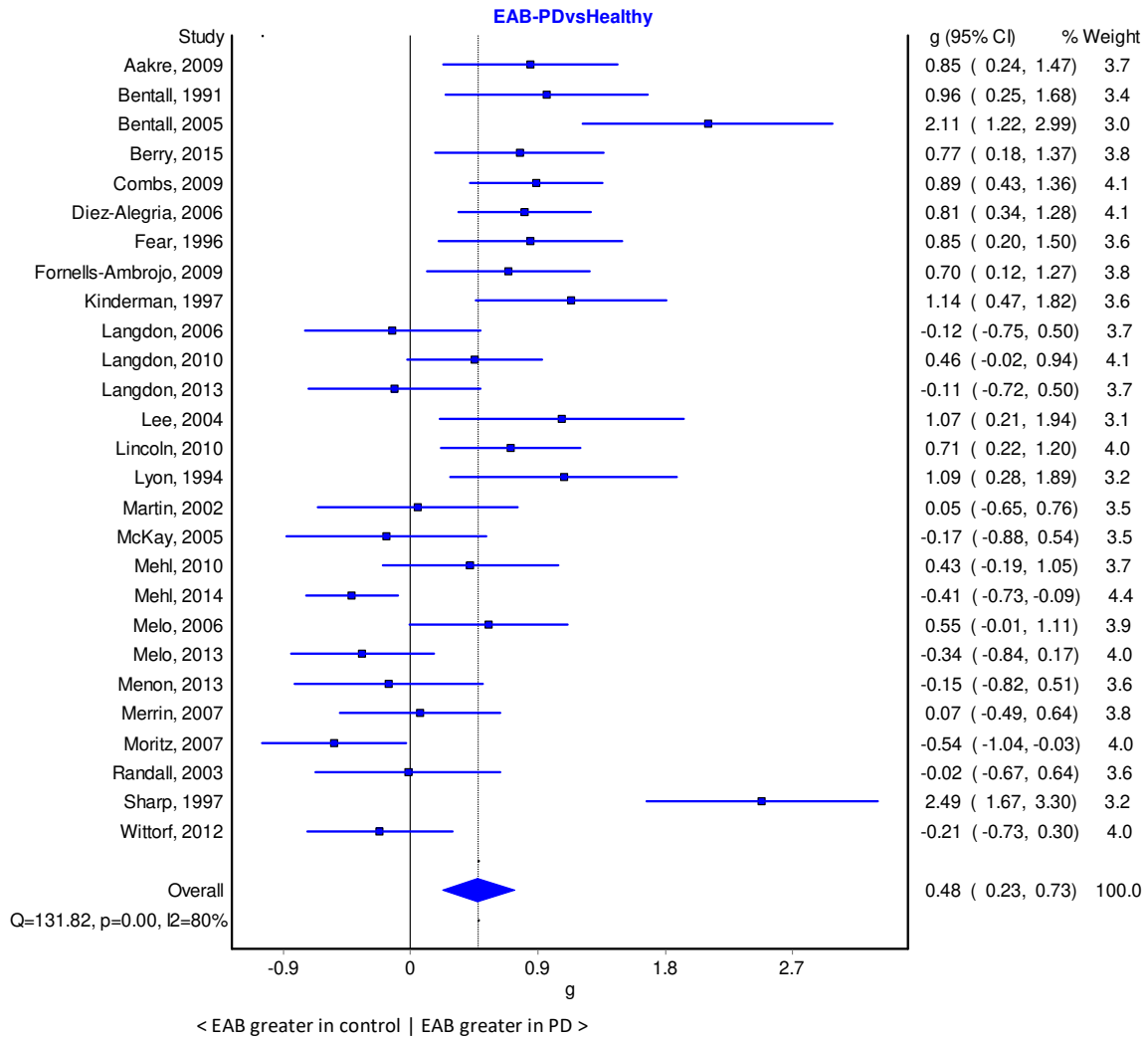
Abbreviations: GDs, grandiose delusions; PDs, persecutory delusions.

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

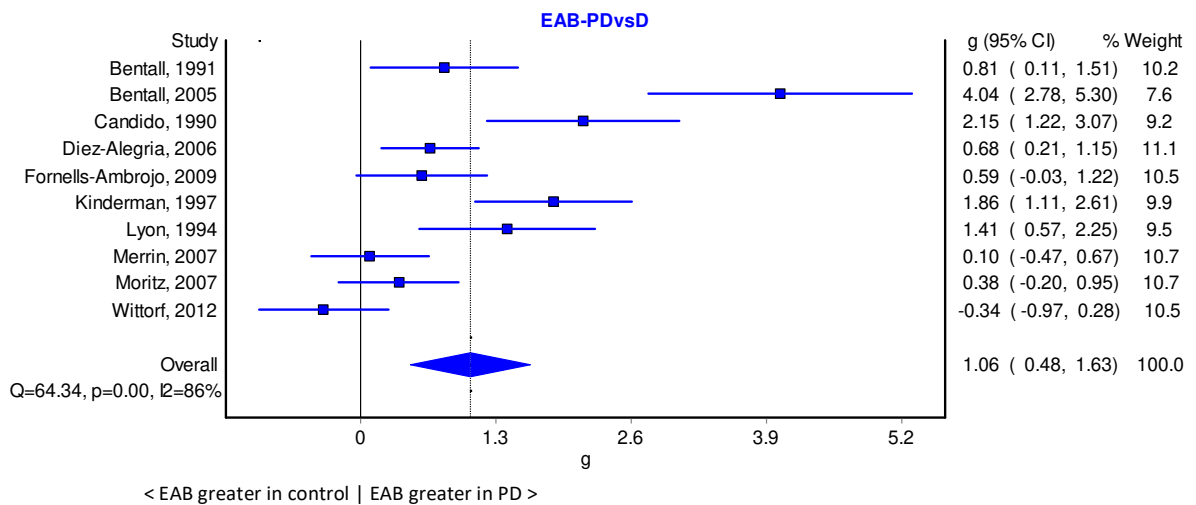
^b'Depression differences' (quantified using the SMD, d) was a continuous moderator. N = number of studies that provided information on depression differences.

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

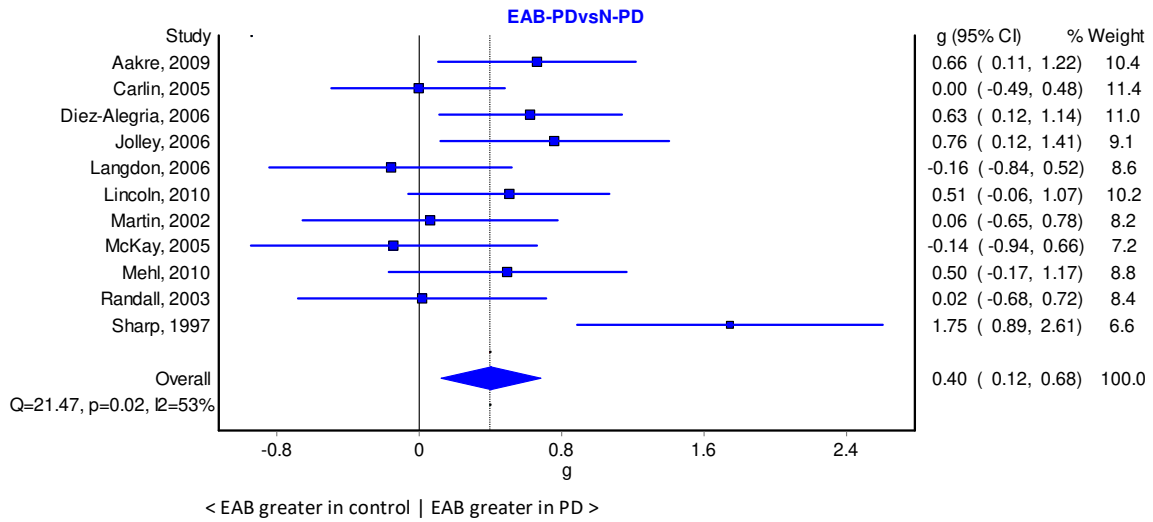
A



B



C



D

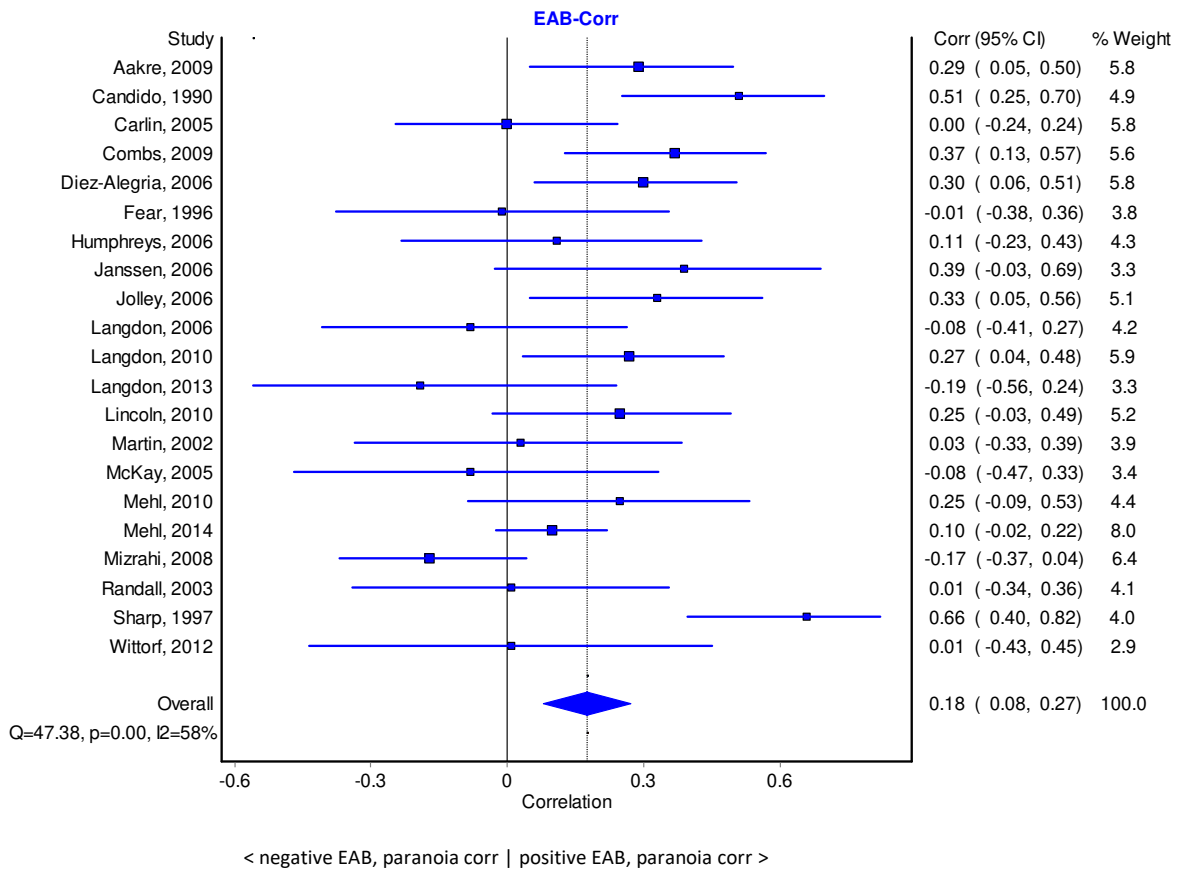


Figure 5. Forest plots for analyses of externalising attributional bias (EAB). (A) Forest plot for comparison of EAB between people with psychosis with persecutory delusions (PDs) and healthy (non-clinical) individuals. (B) Forest plot for comparison of EAB between people with psychosis with PDs and 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.

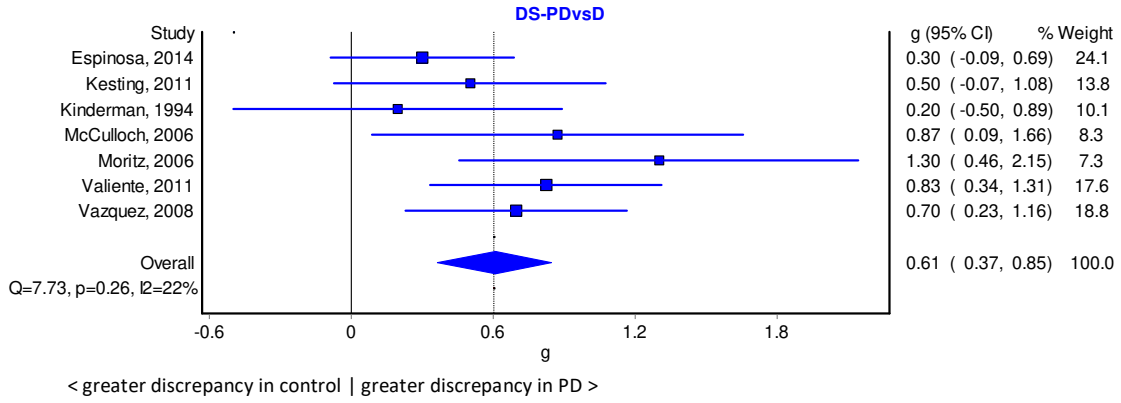


Figure 6. Forest plot for comparison of discrepancy scores^a between people with psychosis with persecutory delusions (PDs) and people with depression.

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

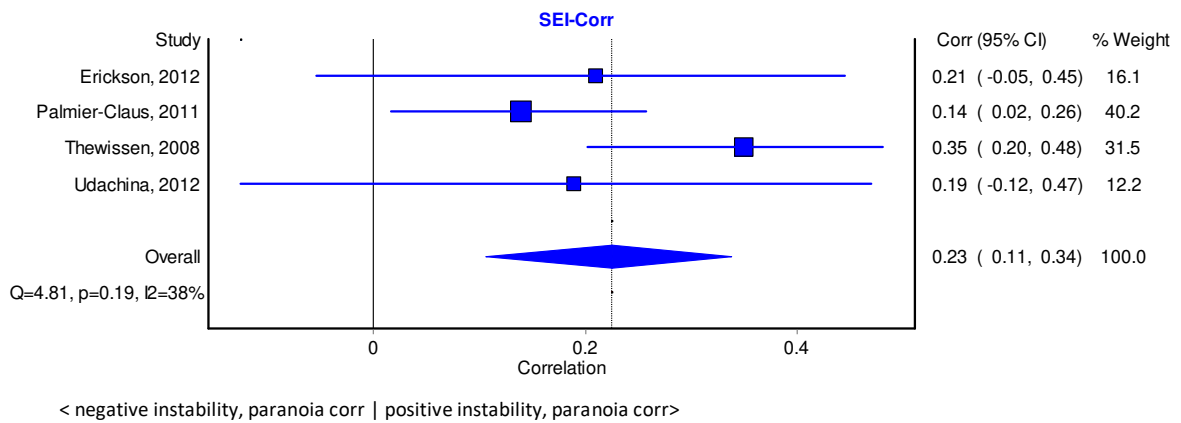


Figure 7. Forest plot of correlation between self-esteem instability and paranoia severity in people with psychosis.

Table 3. Quality of evidence in favour of or against predictions in the ‘paranoia-as-defence’ model

Prediction	Finding
Externalising attributional bias	
People with persecutory delusions will have a significantly greater externalising attributional bias than non-clinical individuals	Moderate quality evidence in favour of defensive prediction (both versions)
People with persecutory delusions will have a significantly greater externalising attributional bias than depressed individuals	Moderate quality evidence in favour of defensive prediction (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 prediction (both versions)
In people with psychosis, the degree of externalising attributional bias will be significantly and positively correlated with paranoia severity	Moderate quality evidence in favour of defensive prediction (both versions)
Explicit self-esteem	
People with persecutory delusions will have a level of explicit self-esteem significantly greater than or similar to non-clinical individuals	High quality evidence against defensive prediction (1994 version)
People with persecutory delusions will have significantly greater explicit self-esteem than depressed individuals	Moderate quality evidence in favour of defensive prediction (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 prediction (1994 version)
In people with psychosis, the degree of explicit self-esteem will be significantly and positively correlated with paranoia severity	High quality evidence against defensive prediction (1994 version)
Implicit self-esteem	

Prediction	Finding
People with persecutory delusions will have significantly lower implicit self-esteem than non-clinical individuals	Low quality evidence in favour of defensive prediction (2001 version)
People with persecutory delusions will have similar implicit self-esteem to depressed individuals	Low quality evidence in favour of defensive prediction (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 prediction (2001 version)
In people with psychosis, the degree of implicit self-esteem will be significantly and negatively correlated with paranoia severity	Low quality evidence against defensive prediction (2001 version)
Self-esteem discrepancy	
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 prediction (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 prediction (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 prediction (2001 version)
In people with psychosis, the degree of implicit-explicit self-esteem discrepancy will be significantly and positively correlated with paranoia severity	Moderate quality evidence against defensive prediction (2001 version)
Self-esteem instability	

Prediction	Finding
In people with psychosis, the degree of self-esteem instability will be significantly and positively correlated with paranoia severity.	High quality evidence in favour of defensive prediction (2001 version)
