This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Facial Affect Recognition in Psychosis

Natalie Sarah Bordon

Submitted in part fulfilment of the degree of Doctorate in Clinical Psychology
The University of Edinburgh
July 2015
Table of Contents

D.Clin.Psychol. Declaration of own work ................................................................. 5
Acknowledgments ................................................................................................. 6
Research Portfolio Abstract .................................................................................... 7
Chapter 1- Systematic Review ............................................................................... 8
Abstract ................................................................................................................. 9
Introduction ......................................................................................................... 10
Method .................................................................................................................. 14
Results ................................................................................................................. 15
Discussion ............................................................................................................ 18
Conclusion .......................................................................................................... 20
References .......................................................................................................... 20
Figure 1. Search Process ....................................................................................... 25
Table 1. Study Characteristics .............................................................................. 26
Table 2. Quality Analysis ..................................................................................... 27
Table 3. Fear Recognition .................................................................................... 28
Table 4. Anger Recognition .................................................................................. 29
Table 5. Disgust Recognition ............................................................................... 30
Table 6. Overall Emotion Recognition .................................................................. 30
Table 7. Intervention Programmes ........................................................................ 31
Appendix 1. Quality Assessment Tool .................................................................. 32
Appendix. 2. AMSTAR Quality Assessment ......................................................... 36
Chapter 2- Systematic Review and Meta-Analysis ................................................. 39
Abstract .............................................................................................................. 40
Introduction ....................................................................................................... 41
Discussion ......................................................................................................... 52
Summary of Findings ........................................................................................... 52
Does FAR cause psychosis? ................................................................................ 53
Research recommendations ................................................................................. 54
Limitations .......................................................................................................... 54
Conclusion and implications ............................................................................... 55
References ......................................................................................................... 56
Figure 1. Search process for non-affective psychosis vs. affective psychosis/bipolar disorder .................................................................63
Figure 2. Search process for active vs non-active psychosis ........................................64
Figure 3. Search process for conversion to psychosis ...............................................................65
Figure 4. Search Process for FRT ..............................................................................................66
Table 1. Study characteristics of non-affective psychosis vs. affective psychosis/bipolar ......66
Table 2. Study characteristics of active vs non-active psychosis ............................................67
Table 3. Study characteristics of conversion to psychosis ..........................................................68
Table 4. Study Characteristics ..................................................................................................70
Table 5. Consistency of FAR and Psychosis- Heterogeneity of included studies ..................73
Figure 5. Specificity of FAR and Psychosis- Comparison FAR ability in non-affective psychosis and affective psychosis/bipolar disorder ........................................73
Figure 6(a). Temporality of FAR and Psychosis- Comparison of CHR with control subjects74
Figure 6(b). Temporality of FAR and Psychosis- Comparison of FAR ability in CHR who convert to psychosis with those who do not convert .........................................................74
Figure 7. Temporality of FAR and Psychosis- Comparison of FAR ability in active and non-active psychosis ............................................................................................................................74
Figure 8(a). Experiment Criterion- The effect of facial affect recognition training on recognition of facial affect at post-intervention ........................................................................75
Figure 8(b). Experiment Criterion- The effect of facial affect recognition training on recognition of facial affect at 1-week follow up ........................................................................75
Figure 8(c). Experiment Criterion- Facial affect recognition training versus active control groups at post-intervention .................................................................75
Figure 9(a). Experiment Criterion- The effect of facial affect recognition training on negative symptoms..........................................................................................................................76
Figure 9(b). Experiment Criterion- The effect of facial affect recognition training on positive symptoms ..........................................................................................................................76
Figure 9(c). Experiment Criterion- The effect of facial affect recognition training on general psychopathology ................................................................................................................76
Table 6. Summary of Outcomes ...............................................................................................77
Table 7(a). Quality Analysis of Non-RCT studies- Non-affective psychosis vs affective psychosis ...........................................................................................................................78
Table 7(b) Quality Analysis of Non-RCT studies- Active vs Non-active Psychosis ..........79
Table 7(c) Quality Analysis of Non-RCT studies- conversion to psychosis .......................79
Table 7(d) Risk of Bias Table for FRT Studies .........................................................................80
Table 8. GRADE Assessment of Outcomes ............................................................................80
D.Clin.Psych. Declaration of own work
This sheet must be filled in (each box ticked to show that the condition has been met), signed and dated, and included with all assessments - work will not be marked unless this is done

Name: Natalie Bordon

Assessed work: Case Study Conceptualisation Research proposal Case Study
(please circle) SSR Essay Question Paper Thesis

Title of work: An Evaluation of a Facial Affect Recognition Training Programme in a Forensic Schizophrenic Population

I confirm that all this work is my own except where indicated, and that I have:

- Read and understood the Plagiarism Rules and Regulations
- Composed and undertaken the work myself
- Clearly referenced/listed all sources as appropriate
- Referenced and put in inverted commas any quoted text of more than three words (from books, web, etc)
- Given the sources of all pictures, data etc. that are not my own
- Not made undue use of essay(s) of any other student(s) either past or present (or where used, this has been referenced appropriately)
- Not sought or used the help of any external professional agencies for the work (or where used, this has been referenced appropriately)
- Not submitted the work for any other degree or professional qualification except as specified
- Acknowledged in appropriate places any help that I have received from others (e.g. fellow students, technicians, statisticians, external sources)
- Complied with other plagiarism criteria specified in the Programme Handbook
- I understand that any false claim for this work will be penalised in accordance with the University regulations

Additionally, for SSR and Thesis submissions:

- Received ethical approval from the University of Edinburgh, School of Health
- OR
- Received ethical approval from an approved external body (e.g. NHS Research Ethics Committee) and registered this application and confirmation of approval with the University of Edinburgh’s School of Health’s ethical committee

Signature ………………… Date …01/07/2015…

Please note:
a) If you need further guidance on plagiarism, you can:
i/ Speak to your director of studies / thesis supervisor
ii/ View university regulations at http://www.ed.ac.uk/schools-departments/academic-services/policies-regulations
b) Unless otherwise stated, referencing should be in the format of the BPS style guide, which is freely available from the BPS web site
Acknowledgments

I would like to thank my academic supervisors Dr Paul Hutton and Dr Suzanne O’Rourke for their guidance and support throughout my research. Thank you to my parents who have continued to support me throughout my career. Their belief in me has been immense and at times been the only thing to keep me going!
Research Portfolio Abstract

Introduction: This thesis aims to examine the association between facial affect recognition difficulties and aggression in patients with psychosis and then explore the causality of facial affect recognition difficulties in psychosis. A systematic review was carried out to explore the links between emotional recognition deficits and aggression and violence in patients with psychosis. An assessment of causality of facial affect recognition difficulties in psychosis was then carried out using the Bradford Hill Criteria.

Methods: For the first systematic review, a systematic search of databases using predefined inclusion criteria returned 6 papers exploring the links between facial affect recognition difficulties and aggression. For the second review, a series of hypotheses were constructed based on the Bradford Hill Criteria. Where there was no up to date meta-analysis available to answer each of these hypotheses, where possible, a meta-analysis was conducted to test the hypothesis.

Results: The first systematic review revealed mixed findings for the association between emotion recognition and aggression with some studies providing support for this association and others reporting no significant correlations. The second paper presented good evidence for a causal role of facial affect recognition difficulties in psychosis. There is evidence for these difficulties presence in the psychosis population and indeed prior to onset of symptoms, and the difficulties do not appear to remediate when symptoms improve. Evidence is also presented to suggest that facial affect recognition problems can be improved through specific facial affect recognition training programmes.

Conclusions: Given the mixed findings of the systematic review, no definitive conclusions can be made regarding the links between emotion recognition deficits and aggression. While there appears to be evidence of facial affect recognition difficulties being one of several possible causes of psychosis and this mechanism is considered in line with current models of psychosis. The results of both reviews are considered in line with assessments of their quality and recommendations for future trials are discussed.
Chapter 1 - Systematic Review

Facial affect recognition and aggression in patients with psychosis: A systematic review

Natalie Bordon*¹, Paul Hutton² and Suzanne O’Rourke²

* Corresponding author email address; Natalie.Bordon@nhs.net
¹ The State Hospital, Lanark, NHS The State Hospital.
² The School of Health in Social Science, the University of Edinburgh.

Prepared for submission to Psychological Review
Abstract
While a correlation between suffering from psychosis and an increased risk of engaging in aggressive behaviours has been established, many factors have been explored which may contribute to increasing this risk. Patients with a diagnosis of psychosis have been shown to have significant difficulties in facial affect recognition (FAR) and some authors have proposed that this may contribute to increasing the risk of displaying aggressive or violent behaviours. A systematic review of the current evidence regarding the links between facial affect recognition and aggression was conducted. Results were varied with some studies providing evidence of a link between emotion recognition difficulties and aggression, while others were unable to establish such an association. Results should be interpreted with some caution as the quality of included studies was poor due to small sample sizes, insufficient power and limited reporting of results. Adequately powered, randomised controlled studies using appropriate blinding procedures and validated measures are therefore required.

Word Count: 5584
Introduction

There is an association between psychosis and aggressive behaviour (Douglas, Guy, & Hart, 2009; Fazel, Gulati, Linsell, Geddes, & Grann, 2009; Mullen, 2006) with many factors being identified as possibly mediating this risk, including positive symptoms (Link, Stueve, & Phelan, 1998; Jeffrey W. Swanson et al., 2006) substance misuse (Eriksson, Romelsjö, Stenbacka, & Tengström, 2011; Fazel et al., 2009; Jeffrey W. Swanson, Holzer III, Ganju, & Jono, 1990) and co-morbid personality disorders (Hodgins, 2008; Nolan, Volavka, Mohr, & Czobor, 1999). More recently paranoid ideation specifically has been shown to be associated with violence (OR= 2.26) (Coid, Ullrich, Bebbington, Fazel, & Keers, 2016).

Numerous studies have shown that patients with psychosis have a significantly reduced ability to recognise emotions in others, and a recent meta-analysis of over 50 studies confirmed that this impairment is large in magnitude (Kohler, Walker, Martin, Healey, & Moberg, 2009). This is important because a reduced ability to recognise emotional expressions in others is associated with an increased risk of engaging in violence (Stephanie T. Harris & Picchioni, 2013; Hoaken, Allaby, & Earle, 2007; Weiss et al., 2006). Recognition of facial expressions is particularly important in social interactions and deficits in emotional recognition have been reported to be associated with poor social functioning (Hall et al., 2004; Mueser et al., 1996) and can lead to inappropriate behaviours, for example aggression in harmless situations (Hoaken et al., 2007). More specifically difficulties in recognising fear expressions have also been shown to be associated with anti-social behaviour and increased incarcerations (Hastings, Tangney, & Stuewig, 2008; Marsh & Blair, 2008).

It has been established that patients with a diagnosis of psychosis can experience bias in the processing of social information (Freeman & Garety, 2014; Garety, Kuipers, Fowler, Freeman, & Bebbington, 2001). There is strong evidence for a ‘jumping to conclusions’ bias where those with psychosis have been shown to make decisions on the basis of less information than those without psychotic mental health diagnoses (Dudley, Taylor, Wickham, & Hutton, 2015). It has also been established that those with a diagnosis of psychosis can have an externalising attributional style, which refers to the tendency to attribute the causes of negative events to external factors (Bentall, Corcoran, Howard, Blackwood, & Kinderman, 2001; Kinderman & Bentall, 1997). Hostile attributional bias can also occur when a person assigns deliberate intent of another in causing the negative event (S. T. Harris, Oakley, & Picchioni, 2014). Previous research has suggested that externalising and hostile biases are linked with an increase in the potential for violence or aggression (R. Edwards & Bond, 2012; S. T. Harris et al., 2014; McNiel, Eisner, & Binder, 2003) as it has been proposed that those who have a hostile attributional bias may be more likely to misinterpret ambiguous situations as involving threat, leading them to respond in an aggressive way (S. T. Harris et al., 2014).

Poor social understanding or theory of mind, have also been established as being associated with psychosis (Bora & Pantelis, 2013; Brüne, 2005; Garety et al., 2001). Theory of mind is defined as understanding and conceptualising the mental state of others and can allow an
individual to understand and attempt to predict the behaviour of another (Baron-Cohen & Wheelwright, 2004). Impairments in theory of mind, may lead to difficulties in correctly interpreting the intentions of others and if these intentions are deemed to be hostile, this may increase the risk of violence (Bo, Abu-Akel, Kongerslev, Haahr, & Simonsen, 2011; Jeffrey W. Swanson et al., 2006). If an individual has difficulties with FAR, this may have an impact on their theory of mind ability as they may not be able to accurately perceive the emotion displayed by another, thus effecting their ability to correctly perceive the other’s mental state.

Some authors have tried to explore the theoretical underpinnings of this link between violence and difficulties in the recognition of emotions in others. A number have proposed that disruption in the amygdala contributes to these difficulties in social cognition (Adolphs & Spezio, 2006; Gumley, Braehler, & Macbeth, 2014; Rosenfeld, Lieberman, & Jaroskog, 2010; Whalen et al., 2013). More specifically they suggest that the dopaminergic/oxytonergic circuitry is particularly responsible for socio-emotional processing (Gumley et al., 2014). Blair’s Integrated Emotion System (IES) model (Blair, 2005) suggests that a dysfunction within the amygdala impairs the ability to recognize distressing emotions in others and due to this, impacts the development of empathy. Expressions of fear are proposed as distress cues that serve to inhibit aggression and violence from others. The amygdala has been identified as a significant area of the brain in processing emotion (LeDoux, 1998) and lesions in this area have been shown to have an impact on the recognition of fearful expressions (Blair, 2003). Functional imaging studies have shown a reduced activation of the amygdala in those diagnosed with psychopathy during emotional memory (Kiehl et al., 2001) and the recognition of fearful facial expressions (Blair, Colledge, Murray, & Mitchell, 2001), suggesting that the amygdala may be a mediating factor in the link between facial affect recognition and violence. The IES model (Blair, 2005) is important for explaining the underlying cognitive mechanisms of the link between FAR and violence, as if a person is inhibited in their recognition of emotion in others, they may be unable to detect the social cue of fear in others, which would alert them to modify their aggressive behaviour.

Blair (2005) builds upon an empathy model (Marshall, Hudson, Jones, & Fernandez, 1995) which proposes that the first stage in the process of empathy is the recognition of the emotion in others and if this does not happen the subsequent stages of perspective taking, emotion replication and response decision cannot take place. Blair (1995) argues that expressions of fear or sadness act as an unconditioned stimuli, producing an aversive response in the person witnessing it to help teach and develop the social valence of one’s actions. It is proposed that those with psychopathy, due to their cognitive deficit, are less able to learn from this experience and therefore this can result in less socially acceptable behaviour which may involve harm or violence against another person.

Crick and Dodge (1994) proposed a social information processing model in which the first steps involve attending to, encoding and interpreting a social situation before then ascertaining the goals of the situation and responding. Lemerise and Arsenio (2000) have expanded this model to include the role of emotion and have highlighted that it can influence each of the stages. They propose that being able to attend to and recognise others’ emotional
cues is a crucial source of information in not only assessing the situation but also influencing how we choose to respond. Therefore those who have difficulty in accurately interpreting emotional cues may have trouble responding in an appropriate manner. While there are many social cues that we attend to in social situations such as voice prosody and body language, FAR also plays a significant role in this process.

While there are many reasons that someone may be aggressive or commit a violent act, it is useful to categorise these incidents in terms of their motives for doing so to help explore if any common underlying triggers exist. Berkowitz (1993) and Geen (2001) define instrumental aggression as “a premeditated means of obtaining some goal other than harming the victim, and being proactive rather than reactive”. This is then compared to hostile or affective aggression which is defined as “being impulsive, thoughtless (i.e. unplanned), driven by anger, having the ultimate motive of harming the target and occurring as a reaction to some perceived provocation” (Anderson & Bushman, 2002). This distinction is of particular relevance when exploring the role of facial affect recognition difficulties in aggression and violence since a misinterpretation of emotion may be processed as a ‘perceived provocation’ and lead to an unplanned or impulsive episode of aggression.

Lemerise and Arsenio (2000) suggest that people who have difficulty in reading others affective signals have problems in responding in an appropriate manner to that person in a given situation. As some researchers have made links between difficulties in recognising specific emotions, namely negative ones (J. Edwards, Pattison, Jackson, & Wales, 2001; Kohler et al., 2003) and also tendencies for neutral emotional expressions to be interpreted as negative (Kohler et al., 2003) it is suggested that this can contribute to escalation in aggression and also a failure to recognise resolution signals should they be present (Silver, Goodman, Knoll, Isakov, & Modai, 2005). While this previous research has suggested that failure to correctly interpret social situations can lead to aggression, none to date has examined the specific type of aggression displayed and its links to facial affect recognition difficulties in order to further support this hypothesis. Given that the models proposed are centred around misinterpretation in the recognition of emotions and inability to identify distress cues, further exploration of the types of aggressive incidents that patients become involved in are required to ascertain if they are caused by a misidentification of emotions in others or are due to a desire to attain a predefined goal in which emotion recognition my not play a part.

Hodgins (2008) proposes that violent offenders should be categorised into one of two groups, ‘early’ or ‘late’ starters. Early starters are defined as engaging in aggression and violence from a younger age and would typically meet diagnostic criteria for conduct disorder or antisocial personality disorder prior to the onset of psychosis. Late starters are those who have a lack of violent behaviour premorbidly and only engage in such behaviour after illness onset. In using this model, Hodgins (2008) found that those classified as early starters had significantly higher rates of aggression and violence compared to late starters. Jeffrey W Swanson et al. (2008) also reported higher levels of violence in those with a history of childhood conduct problems. They found no link between positive symptoms and violence in
the early starters cohort, but did find an association between positive symptoms and violence in the late starters group. Jeffrey W Swanson et al. (2008) therefore concluded that antisocial problems and psychotic symptoms are distinct disorders which both increase the risk of violence in those with psychosis independent of each other.

This distinction between onset of violence is of particular relevance when examining the role of FAR in aggression in patients with psychosis. Studies have demonstrated the link between FAR problems and aggression in those with an anti-social personality disorder (Blair & Cipolotti, 2000; Blair et al., 2004; Marsh & Blair, 2008) and psychopathy (Hastings et al., 2008; Marsh & Blair, 2008). As anti-social personality disorder is around 5-11 times more prevalent in those with psychosis than healthy age and sex matched controls (Tengstrom, Hodgins, Grann, Langstrom, & Kullgren, 2004), these are potential confounding variables when trying to examine the link between FAR and aggression in those with a diagnosis of psychosis, as we cannot be sure that the link exists due to the psychosis rather than the personality disorder. However, this is further complicated by the fact that some authors have suggested that the FAR difficulties can be present in those at clinical high risk of psychosis, prior to onset of any psychosis symptoms (Corcoran et al., 2015; Gibson, Penn, Prinstein, Perkins, & Belger, 2010; Piskulic et al., 2016), therefore violence prior to illness onset may still be driven by psychosis related factors. To help explore the association between FAR and aggression in patients with psychosis, future studies need to examine violence onset and measure and control for confounding factors such as anti-social personality disorder and psychopathy to help ensure that aggression is not due to pre-existing personality traits.

Walz and Benson (1996) reported that aggressive men with cognitive impairment and no history of violence were more likely to label an ambiguous expression as angry compared to their peers. As they also reported that those with cognitive impairment with no history of violence did not demonstrate greater difficulty on emotion labelling and discrimination, they propose that recognition of anger is a mediating factor in aggression. Given the models explained previously (Blair, 2005; Crick & Dodge, 1994; Lemerise & Arsenio, 2000), it may be that difficulties in the recognition of specific emotions may be more of a mediating factor between FAR and aggression rather than global emotion recognition problems. Fear and anger have been reported to act as restraint-producing environmental cues (Weiss et al., 2006) which help to regulate behaviour when noticed. Similarly disgust has also been described as a critical social cue to indicate that something is inappropriate or disliked (Hofer et al., 2009), therefore correct identification of these particular emotions may be more important in the perpetration of violent behaviours rather than an overall difficulty.

Malone, Carroll, and Murphy (2012) provided a wide-ranging theoretical review of the role of FAR in aggression in patients with psychosis. They found only three small studies with directly relevant data, the results of which were inconclusive. Unfortunately Malone et al did not provide effect sizes, and they did not examine study quality in a standardised or systematic way. Since a number of new studies have been published recently, a further review is timely. The aim of this systematic review is therefore to revisit the literature to determine whether a reduced ability to recognise facial affect (specifically anger, fear, disgust
and also overall facial affect recognition ability) is associated with increased aggression in people diagnosed with psychosis. Study quality will be systematically assessed and used to inform interpretation of the observed effect sizes.

Method

Search strategy
A search of the literature was carried out from all years up to March 2016 from the following databases: Medline, Embase, PsychInfo and Web Science. Search terms that were used were ‘facial affect recognition, facial emotion recognition, facial affect recognition training, social cognition, emotion perception, schizophrenia’ and ‘schizoaffective’. This systematic review was completed in accordance with PRISMA and AMSTAR guidelines. Given that there may be studies which have been completed in the area of interest but may not have been published, a search of various trial registers was carried out to identify any unpublished studies. The Cochrane Group Trials Register (CENTRAL), the US government clinical trials register (clinicaltrials.gov), European Union clinical trials register (clinicaltrials-register.eu), World Health Organisation (apps.who.int/trialsearch) and Current Controlled Trials Ltd (controlled-trials.com) were all searched in May 2016 and no unpublished studies that met inclusion criteria were identified. Reference sections within the articles which met the inclusion criteria were searched by hand to identify any further papers.

Inclusion Criteria
Eligible studies had to report usable cross-sectional, longitudinal or experimental data relating to the association between aggression and facial affect recognition in people with a diagnosis of psychosis. As such, a wide range of study designs were eligible for inclusion, including randomised controlled trials, correlational studies and group-comparison studies. Eligible studies had to provide data relating to adult (18 years+) participants with a diagnosis of non-affective psychotic illness (schizophrenia, brief psychotic disorder) or schizoaffective disorder. In order to be included in the final analysis, at least 50% of participants were required to have a diagnosis of non-affective psychosis. Studies which contained participants identified as having meeting diagnostic criteria for anti-social personality disorder or psychopathy were excluded.

Data extraction and Outcomes
Eligible studies had to measure facial affect recognition ability using reliable and valid tools such as the Facial Emotion Identification Test (FEIT) or the Penn Emotion Acuity Test (PEAT), or a less structured measure taken from an established facial expression stimuli tool such as the Pictures of Facial Affect (PFA) (Ekman & Friesen, 1976). Many of these measures are based on Ekman’s pictures of facial affect making their results comparable. The Bell Lysaker Emotion Recognition Scale (BLERT) and Penn Emotion Recognition Task (ER40) have previously been shown to have good psychometric properties for measuring facial affect recognition in patients with a diagnosis of psychosis (Pinkham, Penn, Green, & Harvey, 2015), however it is recognised that further evidence of the psychometric properties of these assessments for use with patients with psychosis is required.
An initial scoping review suggested that there may be a limited number of studies within this area, therefore no a priori decisions were made in relation to measures for violence and aggression. Given the previously discussed models highlighting the roles of fear, anger and disgust recognition in aggression, data was extracted for these individual emotions where available. Where possible results were converted to an effect size (Cohen’s d, with 95% confidence intervals) using Borenstein, Hedges, Higgins, and Rothstein (2009), Cohen (1988) or the Campbell Collaboration effect size calculator (Wilson).

Quality Assessment
The AHRQ (Agency for Healthcare Research and Quality) (Health & Services, 2012) is recommended by the CRD (Centre for Reviews and Dissemination)(2009) as a suitable tool for assessing the quality of observational studies. The CRD further advises that the tool should be adapted and tailored towards the individual requirements of the systematic review. Due to this an adapted version of the AHRQ was utilised and included the domains of selection bias, detection bias, statistical power, validity of measures and method of analysis. Each item within the domains was rated using the tool and assigned a rating of either ‘yes’, ‘no’, ‘partially’, ‘or unclear’.

Results
Study Selection
Figure 1 outlines the process of study selection. After removal of duplicates, the initial search returned 2439 articles, including conference abstracts and dissertations. On reviewing the titles and abstracts, the majority of these papers were excluded as it was clear that they did not examine FAR or aggression. The full text of 35 papers was then reviewed and from this 5 papers met the criteria for inclusion in the study. One further study was identified as relevant from reviewing the reference lists of the included studies bringing the total number of studies in the review to 6. All of these studies were correlational studies.

Characteristics of Included Studies
Of the 6 studies included, 5 studies compared FAR in patients with psychosis who had violent histories to a group of patients with no previous violence. Only 2 of these studies excluded participants with anti-social personality disorder and one study reported excluding participants who scored highly on the Psychopathy Checklist (PCL-R). One study examined FAR ability in patients with a diagnosis of psychosis and a history of violence and the treatability of FAR. This study did not measure or control for anti-social personality disorder. The remaining study did not specifically recruit participants with a diagnosis of psychosis from a forensic service or with a criminal background. All but one study (Demirbuga et al., 2013) recruited participants from inpatient units. The types of measures that were used to assess emotional recognition varied across the studies and included the Penn Emotion Recognition Test (PERT), Japanese and Caucasian Brief Affect Recognition Test and less well defined emotion recognition tasks. Additionally given the variation in measures used, the specific emotions investigated within each of the individual studies also varied as did the
intensity of the emotions presented. In relation to measures of aggression some studies reviewed official criminal records and records of reported aggression in hospital to measure aggression and violence. Two studies also utilised more specific tools to measure aggression (Hostility and Aggression Questionnaire and Life History of Aggression Scale). Full details of study characteristics are included in Table 1.

Results of Quality Assessment
The completed quality assessment of the studies using the AHRQ tool can be found in Table 2. All of the studies adequately reported the demographics of the participants that were included in the study, giving details on their age, gender, ethnicity, intelligence/educational achievement and also the setting that they were recruited from. Most studies also examined official criminal records to confirm the presence of violence and aggression in the patients’ past behaviours, rather than relying solely on self-report measures. It should be noted however that where validated aggression tools were used, the psychometric properties of these tools was not reported. The Life History of Aggression scale (LHA) relies mainly on self-report. Given that aggression is a socially undesirable behaviour, respondents may not wish to give a negative impression and therefore may not disclose the true extent of their aggression (Krahé, 2013). Similarly no descriptions of the psychometric properties of the individual facial affect recognition tools were provided or confirmation that the tools are valid and reliable for assessing facial affect recognition within a psychosis population. While the studies reported the tools which they used, these varied considerably in terms of the number of faces presented, the individual emotions assessed and the intensity of the emotions presented therefore reducing the comparability of the findings.

The main weakness of all of the included studies was that they lacked sufficient power to reliably detect small to moderate effect sizes (Faul, Erdfelder, Lang, & Buchner, 2007) due to their small sample sizes. Related to this was a lack of justification of the sample sizes that were used. A further potential weakness that was unable to be sufficiently assessed across all studies, was the blinding of study personnel to the aggressive history status of the participants that they were assessing. While the majority of studies appear to have used appropriate methods of statistical analysis, not all of the studies have reported sufficient details of the results, for example confidence intervals were missing across half of the studies (Field, 2013). This further impacts the interpretation and comparability of studies. To assist with interpretation of findings, studies have been assigned overall quality ratings. While there was some variability in quality across included studies, given these weaknesses all results should be interpreted with some caution.

Recognition of fear
Four of the studies specifically examined and reported data on the recognition of fear and its links to aggression in psychosis (Table 3). Frommann, Stroth, Brinkmeyer, Wolwer, and Luckhaus (2013) reported that a group of patients within a forensic setting were significantly more impaired on the recognition of fear than a group with no history of violence (d=0.69, CI, 0.03-1.34). However, Wolfmüller et al. (2012) and Demirbuga et al. (2013) found no significant differences between patients from forensic and non-forensic settings in their
recognition of fear. Wolfkühler et al. (2012) found a non-forensic patient group were less able to recognise fear than those in a forensic setting. Bedwell et al. (2013) reported that while patients with psychosis made more errors in the recognition of fear when compared to healthy controls, this difference was not significant. Demirbuga et al. (2013) found that correct responses to the recognition of fear was the lowest across all emotions examined and that it was frequently confused with surprise, and Weiss et al. (2006) found that fearful faces were most commonly misidentified as neutral (41.3%) followed by sad (30.3%).

There was a small association between general criminal behaviour and poor fear recognition in one small study (d=0.26) (Weiss et al., 2006), however no significant associations were found when comparing it to violent arrests or history of aggression (Weiss et al., 2006). However, the number of violent arrests was positively associated with the misinterpretation of the other emotional faces examined (happy, anger, sad and neutral) as fear (d=1.5)(Weiss et al., 2006). While the Frommann et al. (2013) study was of higher quality than the others and a moderate effect size was reported, given the imprecision of the effect size and lack of other studies supporting this relationship, there does not appear to be sufficient evidence to support poor fear recognition resulting in more displays of aggression.

**Recognition of anger**

Four studies reported data on the recognition of anger (Table 4). One study found that both those with a diagnosis of psychosis and a history of violence and those without were impaired on their recognition of anger in comparison to healthy controls (history of violence d=1.61, CI 0.96-2.26; no history of violence d=1.13, CI 0.52-1.74) (Wolfkühler et al., 2012). When a comparison of those within a forensic and non-forensic service's recognition of anger was carried out no significant differences were found (Wolfkühler et al., 2012). This finding was further replicated by Frommann et al. (2013) and Demirbuga et al. (2013) who also found that while a group of patients with a history of violence performed poorer on anger recognition compared to those without, this difference was not significant.

One study found anger was most commonly misinterpreted as neutral affect (36.9%) followed by sadness (29.2%)(Weiss et al., 2006). Number of general arrests was shown to be associated with incorrectly identifying other emotions as angry (d=0.33), however the Life History of Aggression scale was found to be negatively associated with the misinterpretation of faces as angry (d=0.35)(Weiss et al., 2006). Weiss et al. (2006) also reported that while the patients with a history of violence were more likely to misinterpret anger as sadness, those without a history of violence were more likely to misinterpret anger as a neutral affect. While there may be association between an impaired ability to recognise anger and higher numbers of general arrests (d=-0.15), there was no association with a history of violence (Weiss et al., 2006).

**Recognition of disgust**

Three studies provided details on the recognition of disgust (Table 5). Wolfkühler et al. (2012) reported that the patients with a history of violence’s recognition of disgust was significantly greater than the non-violence history patient group (d=0.74, CI, 0.22-1.27) and
that this result remained significant when controlling for differences in the cognitive, excitement and depression and anxiety factors on the PANSS between the two groups. This differs from the results reported in Frommann et al. (2013) which found no significant differences between them. While the effect size found was large, given the spread of the confidence interval reduces the precision of this finding and as the Wolfkühler et al. (2012) study was of a lower quality than the Frommann et al. (2013) one, it is difficult to form any significant conclusions. Demirbuga et al. (2013) found that 30% of patients with psychosis misinterpreted a disgusted face as an angry face, although this study was of low quality.

**General emotional recognition**

Four studies provided data on overall emotional recognition scores (Table 6). Wolfkühler et al. (2012) found both those within a forensic and non-forensic setting were significantly impaired relative to healthy controls. The patients with a history of violence performed slightly better than those without, but the difference was not significant. When controlling for the cognitive, excitement and depression and anxiety factors on the PANSS, they reported that those in the forensic setting significantly outperformed the non-forensic group in general emotion recognition. Silver et al. (2005) also found patients were significantly impaired relative to healthy controls on their ability to recognise emotions. They also reported a significant difference in the patients with a history of violence and those with no previous violence’s ability to recognise emotions with those who engaged in violence previously demonstrating better emotional recognition (d=2.07, CI, 1.49, 2.65). This contrasts with Frommann et al. (2013) who found that patient in a forensic setting were significantly poorer in general emotional recognition than a non-forensic setting (d=0.92, CI, 0.26, 1.59). As both of these studies were assessed as similar quality and are each reporting large effect sizes, it is difficult to draw a definitive conclusion regarding this outcome.

A small association between a higher number of arrests in general and poorer FAR ability was also reported in one study (d=-0.10), however no association between number of violent arrests or higher scores on the Life History of Aggression scale was established (Weiss et al., 2006).

**Intervention programmes**

One study examined an intervention programme designed to improve FAR (Table 7). Combs et al. (2007) carried out a treatment study using a Social Cognition and Interaction Training (SCIT) programme with patients in a forensic setting and found that after treatment compared to controls, the treatment group had not only improved on FAR (d=1.34) but also reduced the number of aggressive incidents on the ward (d=1.57). As they used PANSS change scores to control for psychopathology in their analysis, they concluded that improving FAR must play an important role in aggression given the concurrent reduction.

**Discussion**

Overall results of emotional recognition performance between people with a diagnosis of psychosis who are and are not violent are mixed, with some studies of similar quality reporting large differences in the direction of non-violent patients performing better
(Frommann et al., 2013) and others demonstrating large effects in the opposite direction (Silver et al., 2005; Wolfkühler et al., 2012). While some studies focused on the correct identification of emotions, others explored these errors in more detail and provided data on misinterpretation of specific emotions including misidentification of anger as neutral affect and sadness (Weiss et al., 2006). While these studies had some weaknesses in relation to quality, given the results there appears to be some evidence of the importance of the misinterpretation of fear in relation to aggression and violence (Weiss et al., 2006) and this is an area which requires further investigation with more robust studies. Further exploration of these types of errors would help provide evidence for a social information processing model of aggression in psychosis, because if patients are misinterpreting emotions then their decision making in their choice of response will be negatively affected and may lead to inappropriate behaviours such as violence.

Although deficits in the recognition of fear have been proposed to be associated with more aggressive and violent incidents (Blair, 2003, 2005; Blair et al., 2001), the findings from this review are also somewhat varied. While one study appeared to support this theory (Frommann et al., 2013) with evidence of a moderate effect size that patients with a history of violence are more impaired on fear recognition than patients without a history of violence, other studies did not support this finding (Demirbuga et al., 2013; Weiss et al., 2006; Wolfkühler et al., 2012). Rather than ability to recognise fear in others, what appeared to be more significant was the misinterpretation of fear in relation to an association with violence and aggression (Weiss et al., 2006). When examining anger recognition there appears to be some evidence that patients with psychosis are in general more impaired on this than controls, however there were no significant differences found between those with a history of violence and those without (Demirbuga et al., 2013; Frommann et al., 2013; Wolfkühler et al., 2012). There appeared to be a small negative association between misinterpretation of other emotions as angry and history of aggression. Once again, given the diverse results and small effect sizes, no definitive conclusions can be made regarding the role of anger recognition in violent incidents.

**Limitations**

While some studies have produced large effect sizes these must be interpreted with caution as all of the studies in this review involved small sample sizes and varied in the specific emotions and intensity of emotions which they investigated. Furthermore given the variation of emotions examined and sample populations, direct comparisons between all included studies was not possible. While some studies recruited patients from forensic mental health services (Demirbuga et al., 2013; Frommann et al., 2013; Luckhaus, Frommann, Stroth, Brinkmeyer, & Wolwer, 2013; Wolfkühler et al., 2012) others recruited patients with a diagnosis of psychosis and subsequently measured levels of aggression (Weiss et al., 2006), therefore it is likely that the violent acts committed by those within forensic services may be more serious and involve greater harm to the victim.

As discussed earlier, none of these studies differentiated between those participants who perpetrated violence prior to illness onset or explored if these violent acts were premeditated
or more impulsive and in response to misunderstanding a situation which may have provided further evidence in the support of the social information processing model (Crick & Dodge, 1994; Lemerise & Arsenio, 2000). Additionally not all of the studies measured and controlled for confounding variables such as personality disorder which has been shown to impact facial affect recognition. Given the somewhat conflicting results and lack of control of confounding variables further high quality studies are required in this area to help establish the links between facial affect recognition and aggression. Establishing an association between aggression and FAR difficulties is particularly important in not only understanding why aggression may occur in patients with psychosis but will also have clinical implications in the treatments that patients with a diagnosis of psychosis who perpetrate violence are recommended to receive in order to reduce their risk of future violence and aggression as well as improving their social functioning.

The methodological quality of this systematic review was self-assessed using AMSTAR (Shea et al., 2007) quality assessment criteria. This tool consists of 11 concise criterion items and a score of 1 is given if the criterion is met and a score of 0 if it is not met, unclear or not applicable. According to this assessment, the main limitation of the current review was non-duplication of study selection and data extraction by two independent assessors. This was a consequence of limited resources available to conduct the review, however, the overall score of 8 suggests the review was otherwise high quality (Sharif, Janjua-Sharif, Ali, & Ahmed, 2013).

**Conclusion**

While difficulties in FAR are established in patients with psychosis (Kohler et al., 2009) and models such as the Integrated Emotional System (Blair, 2005) and social information processing model (Crick & Dodge, 1994) have been proposed to explain the role of these difficulties in aggression and violence, a definitive evidence base for this association is somewhat lacking. As there is established evidence that people with a diagnosis of psychosis can experience a range of reasoning and attributional biases, it may be that FAR difficulties contribute to these established biases, and mediate the incorrect interpretation of social situations. If an individual has a hostile attributional bias and also difficulties with FAR, it may be misinterpretation of others emotional states active this bias and may increase the risk of aggression. Given the varied results for the links between aggression and both general emotion recognition and also for the specific emotions thought to be important in mediating the risks of aggression, no formal conclusions can be made for the current evidence base about the causal role of FAR in aggression in individuals with a diagnosis of psychosis.

**References**


http://schizophreniabulletin.oxfordjournals.org/content/early/2015/10/31/schbul.sbv150.abstract


Freeman, D., & Garety, P. (2014). Advances in understanding and treating persecutory delusions: a review. *Social psychiatry and psychiatric Epidemiology, 49*(8), 1179-1189. doi:10.1007/s00127-014-0928-7


Figure 1. Search Process

Number of record identified through database searching: 4350

Number of records after duplicates removed: 2835

Number of records screened (abstract/description): 318

Number of records excluded on basis of title: 2517

Number of full text reports screened for eligibility: 35

Number of records excluded on abstract: 283

Number of full text reports excluded: 30

- 24 Makes no links between FAR and aggression
- 2 Non-clinical population
- 1 Focuses on impact of psychopathy on FAR
- 3 systematic reviews

Number of studies identified from reference section for inclusion: 1

Number of studies included in the review: 6
Table 1. Study Characteristics

<table>
<thead>
<tr>
<th>Trial</th>
<th>Design</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Location</th>
<th>FAR measures</th>
<th>Aggression measure</th>
<th>Exclusion of PD</th>
<th>Inpatients</th>
<th>Age (years) mean (s.d)</th>
<th>Female n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fromman, et al. (2013)</td>
<td>Group comparison, observational</td>
<td>19 forensic patients*</td>
<td>19 non-violent patients</td>
<td>Germany</td>
<td>Pictures of Facial Affect (PFA)</td>
<td>index offence records</td>
<td>Anti-Social PD and Psychopathy excluded</td>
<td>100%</td>
<td>Violent 35.3 (8.2)</td>
<td>NonViolent 34.7 (10.6)</td>
</tr>
<tr>
<td>Combs, et al. (2007)</td>
<td>Group comparison, non-randomised trial</td>
<td>18 forensic patients receiving SCIT*</td>
<td>10 forensic patients receiving coping skills*</td>
<td>USA (Oklahoma)</td>
<td>emotion recognition test</td>
<td>Hostility and Aggression questionnaire; aggressive incidents on ward</td>
<td>Not known</td>
<td>100%</td>
<td>Violent 41.3 (11.2)</td>
<td>NonViolent 44.0 (10.6)</td>
</tr>
<tr>
<td>Silver, et al. (2005)</td>
<td>Group comparison, observational</td>
<td>35 violent patients*</td>
<td>35 non-violent patients*</td>
<td>Israel</td>
<td>Penn Emotion Recognition Test (PERT)</td>
<td>hospital records or aggressive incidents</td>
<td>Not known</td>
<td>100%</td>
<td>37.03 (10.30)</td>
<td>0%</td>
</tr>
<tr>
<td>Wolfkuhler, et al. (2012)</td>
<td>Group comparison, observational</td>
<td>30 forensic patients*</td>
<td>30 non-violent patients*</td>
<td>Germany</td>
<td>Japanese and Caucasian Brief Affect Recognition Test</td>
<td>criminal records</td>
<td>Not known</td>
<td>100%</td>
<td>Violent 32.2 (7.3)</td>
<td>NonViolent 32.4 (9.3)</td>
</tr>
<tr>
<td>Weiss, et al. (2006)</td>
<td>Within-group, observational</td>
<td>38 patients*</td>
<td>-</td>
<td>USA (New York)</td>
<td>Penn Emotion Recognition Test (PERT)</td>
<td>Life History of Aggression Scale; number of violent arrests</td>
<td>Not known</td>
<td>100%</td>
<td>37.03 (10.30)</td>
<td>0%</td>
</tr>
<tr>
<td>Demirbuga, et al. (2013)</td>
<td>Group comparison, observational</td>
<td>41 forensic patients*</td>
<td>35 non-violent patients*</td>
<td>Istanbul</td>
<td>emotion recognition test</td>
<td>interview (self-report); review of criminal records</td>
<td>Anti-Social PD excluded</td>
<td>0%</td>
<td>Violent 41.50 (7.56)</td>
<td>NonViolent 39.94(6.79)</td>
</tr>
</tbody>
</table>

*All had received a diagnosis of psychosis*
### Table 2. Quality Analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Unbiased selection of cohort</th>
<th>Selection minimises baseline differences in prognostic factors</th>
<th>Sample size justification report</th>
<th>Sufficient power</th>
<th>Adequate description of the cohort</th>
<th>Validated method for ascertaining aggression/violence history</th>
<th>Validated method for measuring facial affect recognition</th>
<th>Outcome assessment blind to exposure</th>
<th>Analytic methods appropriate</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combs (2007)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Yes</td>
<td>Partially</td>
<td>10/18 Higher</td>
</tr>
<tr>
<td>Demirbuga (2013)</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
<td>7/18 Lower</td>
</tr>
<tr>
<td>Fromman (2013)</td>
<td>Partially</td>
<td>Yes</td>
<td>Partially</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
<td>10/18 Higher</td>
</tr>
<tr>
<td>Silver (2005)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Yes</td>
<td>10/18 Higher</td>
</tr>
<tr>
<td>Weiss (2006)</td>
<td>Partially</td>
<td>NA</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Yes</td>
<td>8/18 Medium</td>
</tr>
<tr>
<td>Wolfkuhler (2012)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Yes</td>
<td>9/18 Medium</td>
</tr>
</tbody>
</table>

To help order the quality of studies 2 points given for ‘yes’ response, 1 point for ‘partially’ and 0 for ‘no’ or ‘can’t tell’. Quality ratings of ‘higher’ given to 10/18, ‘medium’ given to 8 or 9/18 and ‘lower’ given to 6 or 7/18. These ratings are given to help with comparison of studies within the review.
<table>
<thead>
<tr>
<th>Study</th>
<th>Finding</th>
<th>Mean (sd)</th>
<th>Results reported in paper</th>
<th>Effect size</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fromman et al. (2013)</td>
<td>V more impaired on recognition of fear than NV</td>
<td>V= 3.4(2.3)</td>
<td>F=4.447, df= 1, p=0.041</td>
<td>d=0.69*, CI 0.03-1.34</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=5.4(3.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolfkuhler et al. (2012)</td>
<td>NV more impaired on recognition of fear than V</td>
<td>V=2.3(2.2)</td>
<td></td>
<td>d=0.37, CI -0.14-0.88</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=1.6(1.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiss et al. (2006)</td>
<td>General criminal behaviour associated with poor fear recognition</td>
<td>OR=0.624, p=0.0119, CI 0.432-0.901</td>
<td>d=0.26*</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Violent arrests not associated with fear recognition</td>
<td>OR=0.825, p=0.1350, CI 0.641-1.061</td>
<td>d=0.11</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td></td>
<td>History of aggression not associated with fear recognition</td>
<td>OR=0.929, p=0.7468, CI 0.594-1.451</td>
<td>d=0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fearful faces most commonly misidentified as neutral then sad</td>
<td>Fear identified as neutral 41.3% Fear identified as sad 30.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>number of violent arrests was associated with the misinterpretation of faces as fear</td>
<td>OR=1.31, p=0.0573</td>
<td>d=1.5*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demirbuga et al. (2013)</td>
<td>No difference between V and NV on recognition of fear</td>
<td>V=2.6(2.2)</td>
<td></td>
<td>d=0.32, CI -0.78-0.13</td>
<td>Lower</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=3.3(2.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

V= people diagnosed with psychosis who were violent or receiving forensic care ; NV=people diagnosed with psychosis who were not judged to be violent; Sz= People diagnosed with psychosis; HC=healthy controls
*indicates significant result
<table>
<thead>
<tr>
<th>Study</th>
<th>Finding</th>
<th>Mean (sd)</th>
<th>Results reported in paper</th>
<th>Effect size</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fromman et al. (2013)</td>
<td>V more impaired on recognition of anger than NV</td>
<td>V=7.4(1.6) NV=6.5(2.3)</td>
<td>-</td>
<td>d=0.45*, CI, 0.19-1.10</td>
<td>Higher</td>
</tr>
<tr>
<td>Wolfkuhler et al. (2012)</td>
<td>V more impaired on recognition of anger than NV</td>
<td>V=2.9(2.2) NV=3.7(2.4)</td>
<td>-</td>
<td>d=0.35, CI, -0.86-0.16</td>
<td>Medium</td>
</tr>
<tr>
<td>Weiss et al. (2006)</td>
<td>General criminal behaviour associated with poor anger recognition</td>
<td>OR=0.756, p=0.0327 CI 0.585-0.977</td>
<td>-</td>
<td>d=0.15*</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Violent arrests not associated with anger recognition</td>
<td>OR=0.823, p=0.2077 CI 0.607-1.114</td>
<td>-</td>
<td>d=0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>History of aggression not associated with fear recognition</td>
<td>OR=1.093, p=0.4504 CI 0.866-1.379</td>
<td>-</td>
<td>d=0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Angry faces most commonly misinterpreted as neutral then sad</td>
<td>Angry identified as neutral 39.9%</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>History of aggression negatively associated with misinterpretation of</td>
<td>OR=0.53, p=0.0018</td>
<td>-</td>
<td>d=0.35*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>faces as angry</td>
<td>Angry identified as sad 29.2%</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demirbuga et al. (2013)</td>
<td>No difference between V and NV recognition of anger</td>
<td>V=6.1(1.8) NV=6.3(1.8)</td>
<td>-</td>
<td>d=0.11, CI -0.56-0.34</td>
<td>Lower</td>
</tr>
<tr>
<td></td>
<td>Positive correlation between accurate responses to angry faces and</td>
<td>r=0.347, p=0.026*</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>general psychopathology score in violent patients</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

V= people diagnosed with psychosis who were violent or receiving forensic care; NV=people diagnosed with psychosis who were not judged to be violent; Sz= People diagnosed with psychosis; HC=healthy controls

*indicates significant result
### Table 5. Disgust Recognition

<table>
<thead>
<tr>
<th>Study</th>
<th>Finding</th>
<th>Mean (sd)</th>
<th>Results reported in paper</th>
<th>Effect size</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fromman et al. (2013)</td>
<td>V more impaired on recognition of disgust than NV</td>
<td>V=4.7(2.5)</td>
<td>Results reported in paper</td>
<td>d=0.12, CI-0.52-0.76</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=5(2.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolfkuhler et al. (2012)</td>
<td>NV more impaired on recognition of disgust than V</td>
<td>V=2.8(1.9)</td>
<td>t=3.225, df=58, p=0.002</td>
<td>d=0.82*, CI 0.34-1.31</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=1.4(1.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demirbuga et al. (2013)</td>
<td>No difference between V and NV recognition of disgust</td>
<td>V=5.6(2.4)</td>
<td></td>
<td>d=0, CI -0.45-0.45</td>
<td>Lower</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=5.6(2.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30% of all schizophrenic patients misinterpreted disgust as angry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

V= people diagnosed with psychosis who were violent or receiving forensic care; NV=people diagnosed with psychosis who were not judged to be violent; *indicates significant result

### Table 6. Overall Emotion Recognition

<table>
<thead>
<tr>
<th>Study</th>
<th>Finding</th>
<th>Mean (sd)</th>
<th>Results reported in paper</th>
<th>Effect size</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fromman et al. (2013)</td>
<td>V more impaired than NV on emotion recognition</td>
<td>V=41.3(5.2)</td>
<td>t=2.85, df=36, p=0.007</td>
<td>d=0.92*, CI 0.26-1.59</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=47.4(6.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silver et al. (2005)</td>
<td>NV more impaired than V on emotion recognition</td>
<td>V=57.3(4.7)</td>
<td></td>
<td>d=2.07*, CI 1.49-2.65</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=45.33(6.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V more impaired than NV on distinguishing intensity of emotions</td>
<td>V=38.2(5.3)</td>
<td></td>
<td>d=2.57*, CI 1.93-3.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=51.8(5.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolfkuhler et al. (2012)</td>
<td>NV more impaired than V on emotion recognition</td>
<td>V=24.1(8.5)</td>
<td>Controlling for PANSS</td>
<td>d=0.26, CI -0.21-0.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NV=22.0(7.8)</td>
<td>F=4.960, df=1, p=0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiss et al. (2006)</td>
<td>More general arrests associated with poorer emotion recognition</td>
<td>OR=0.828, p=0.0140</td>
<td></td>
<td>d=0.10*</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CI 0.712-0.962</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No association between violent arrests and poorer emotion recognition</td>
<td>OR=0.869, p=0.2111</td>
<td></td>
<td>d=0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CI 0.699-1.082</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No association between history of aggression and poorer emotion</td>
<td>OR=1.013, p=0.8872</td>
<td></td>
<td>d=0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>recognition</td>
<td>CI 0.851-1.205</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

V= people diagnosed with psychosis who were violent or receiving forensic care; NV=people diagnosed with psychosis who were not judged to be violent; Sz= People diagnosed with psychosis; HC=healthy controls *indicates significant result
<table>
<thead>
<tr>
<th>Study</th>
<th>Finding</th>
<th>Mean (sd)</th>
<th>Results reported</th>
<th>Effect size</th>
<th>Quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combs et al. (2007)</td>
<td>SCIT treatment improved emotion recognition (compared to coping skills group)</td>
<td>Pre=11.5(2.6) Post=15.9(1.5)</td>
<td>n²p=0.31</td>
<td>d=1.34*</td>
<td>Higher</td>
</tr>
<tr>
<td></td>
<td>SCIT treatment reduced number of aggressive incidents on the ward</td>
<td>Pre=2.9(2.0) Post=1.0(1.3)</td>
<td>n²p=0.38</td>
<td>d=1.57*</td>
<td></td>
</tr>
</tbody>
</table>

V=violent/forensic schizophrenic population; NV=non-violent schizophrenic population; SCIT=Social Cognitive and Interaction Training; TAR=Training of Affect Recognition

*indicates significant result
Appendix 1. Quality Assessment Tool.

Quality assessment of observational studies


General instructions: Grade each criterion as “Yes,” “No,” “Partially,” or “Can’t tell.” Factors to consider when making an assessment are listed under each criterion.

1. Unbiased selection of the cohort?

<table>
<thead>
<tr>
<th></th>
<th>The participants in the study are likely to be representative of the target population. The recruitment strategy is clearly described and less likely to introduce bias.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>ünst</td>
</tr>
<tr>
<td>No</td>
<td>The sample is not likely to be representative of the target population. The recruitment strategy is not described and/or is likely to introduce bias.</td>
</tr>
<tr>
<td>Partially</td>
<td>The participants are less likely to be representative of the target population. The recruitment strategy is somewhat likely to introduce bias.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

2. Selection minimizes baseline differences in prognostic factors

<table>
<thead>
<tr>
<th></th>
<th>The selection of a comparison group was appropriate and the group are unlikely to differ on factors related to the outcome (besides antisocial factors). Or the authors indicated that 80-100% of confounders (age, sex, education, IQ, ethnicity) were controlled for in the design (matching) or in the analysis (propensity scores).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>ünst</td>
</tr>
<tr>
<td>No</td>
<td>There were clear differences in confounding variables between groups of which &lt;60% were controlled for in the design or analysis.</td>
</tr>
<tr>
<td>Partially</td>
<td>The group differed on confounding variables and/or some (60-79%) of which were controlled for in the design or analysis.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>
3. Sample size justification reported

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Power calculation and effect size estimation was clearly reported.</td>
</tr>
<tr>
<td>No</td>
<td>No evidence or justification of sample size.</td>
</tr>
<tr>
<td>Partially</td>
<td>Limited evidence or justification of sample size.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

4. Sufficient power

G*Power 3.1.6 (Faul, Erdfelder, Lang & Bucher, 2007) was used to calculate sample sizes required for sufficient power. For correlational analyses it is necessary to recruit 21 participants to detect a large effect size \((r=0.5)\), 62 to detect a moderate effect size \((r=0.3)\) and 614 participants to detect a small effect size \((r=0.1)\) with the statistical power of 0.8 at an alpha level of 0.05. For differences between groups it is necessary to recruit 21 to each group to detect a large effect size \((d=0.8)\), 51 to detect a moderate effect size \((d=0.5)\) and 310 in each group to detect a small effect size \((d=0.2)\) with the statistical power of 0.8 at an alpha level of 0.05.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The study has a sample size large enough to detect small to moderate group differences ((d=0.2-0.5)) or correlations ((r=0.1-0.3)) with the statistical power of 0.8 at an alpha level of 0.05.</td>
</tr>
<tr>
<td>No</td>
<td>The study has a sample size large enough to detect large to very large differences or correlations with the statistical power of 0.8 at an alpha level of 0.05.</td>
</tr>
<tr>
<td>Partially</td>
<td>The study has a sample size large enough to detect moderate to large group differences ((d=0.5-0.8)) or correlations ((r=0.3-0.5)) with the statistical power of 0.8 at an alpha level of 0.05.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>
5. Adequate description of the cohort?

<table>
<thead>
<tr>
<th>Yes</th>
<th>The cohort is clearly (&gt;4) specified and defined in terms of baseline demographics (age, gender, ethnicity, setting, IQ)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The sample is poorly described in terms of key baseline demographics (&lt;2).</td>
</tr>
<tr>
<td>Partially</td>
<td>The cohort is less well (&lt;3) specified and defined in terms of baseline demographics (age, gender, ethnicity, setting, IQ/educational achievement)?</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

6. Validated method for ascertaining aggression/violence history?

<table>
<thead>
<tr>
<th>Yes</th>
<th>The psychometric properties of the outcome measure are clearly reported and are valid and reliable in the study population. Official criminal records have been reviewed.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The outcome measure has not been described in any detail and/or has not undergone psychometric evaluation. Self-reported history of aggression only.</td>
</tr>
<tr>
<td>Partially</td>
<td>The outcome measure is described less clearly and psychometric properties have not been described and/or the measure has not been validated in this population. Official records reviewed.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

7. Validated method for measuring facial affect recognition deficits?

<table>
<thead>
<tr>
<th>Yes</th>
<th>The psychometric properties of the outcome measure are clearly reported and are valid and reliable in the study population.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The outcome measure has not been described in any detail and/or has not undergone psychometric evaluation.</td>
</tr>
<tr>
<td>Partially</td>
<td>The outcome measure is described less clearly and psychometric properties have not been described and/or the measure has not been validated in this population.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>
8. Outcome assessment blind to exposure?

<table>
<thead>
<tr>
<th>Yes</th>
<th>The study investigators who assessed outcomes were blind to the aggressive history status of the participants. Participants were blind to the research question.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The study investigators who assessed outcomes were not blind to the aggressive history status of the participants. The participants were not blind to the research question.</td>
</tr>
<tr>
<td>Partially</td>
<td>Either the study investigators who assessed outcomes were blind to the aggressive history status of the participants or the participants were blind to the research question.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

9. Analytic methods appropriate?\(^3\)

<table>
<thead>
<tr>
<th>Yes</th>
<th>The method of statistical analysis was appropriate to the research question being asked. Confidence intervals, p-values and effect sizes are reported.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The method of analysis is not appropriate to the research question and does not provide meaningful results. The study investigators who assessed outcomes were not blind to the antisocial status of the participants. The participants were not blind to the research question.</td>
</tr>
<tr>
<td>Partially</td>
<td>The analysis is appropriate however the findings are not reported in sufficient detail.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

---

\(^3\) As recommended in Discovering Statistics Using IBM SPSS Statistics (Field, 2013).
Appendix. 2. AMSTAR Quality Assessment

AMSTAR – a measurement tool to assess the methodological quality of systematic reviews.

1. Was an ‘a priori’ design provided?
The research question and inclusion criteria should be established before the conduct of the review.
Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a “yes.”
✔ Yes
☐ No
☐ Can't answer
☐ Not applicable

2. Was there duplicate study selection and data extraction?
There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.
Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other’s work.
☐ Yes
✔ No
☐ Can't answer
☐ Not applicable

3. Was a comprehensive literature search performed?
At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.
Note: If at least 2 sources + one supplementary strategy used, select “yes” (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).
✔ Yes
☐ No
☐ Can't answer
☐ Not applicable

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.
Note: If review indicates that there was a search for “grey literature” or “unpublished literature,” indicate “yes.” SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.
✔ Yes
☐ No
☐ Can't answer
☐ Not applicable

5. Was a list of studies (included and excluded) provided?
A list of included and excluded studies should be provided.
Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select “no.”
☐ Yes
✓ No
☐ Can't answer
☐ Not applicable

6. Were the characteristics of the included studies provided?
In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.
Note: Acceptable if not in table format as long as they are described as above.
✓ Yes
☐ No
☐ Can't answer
☐ Not applicable

7. Was the scientific quality of the included studies assessed and documented?
'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.
Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study (“low” or “high” is fine, as long as it is clear which studies scored “low” and which scored “high”; a summary score/range for all studies is not acceptable).
✓ Yes
☐ No
☐ Can't answer
☐ Not applicable

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.
Note: Might say something such as “the results should be interpreted with caution due to poor quality of included studies.” Cannot score “yes” for this question if scored “no” for question 7.
✓ Yes
☐ No
☐ Can't answer
☐ Not applicable

9. Were the methods used to combine the findings of studies appropriate?
For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I²). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).
Note: Indicate “yes” if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.
✓ Yes
☐ No
☐ Can't answer
☐ Not applicable
10. Was the likelihood of publication bias assessed?
An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).
Note: If no test values or funnel plot included, score “no”. Score “yes” if mentions that publication bias could not be assessed because there were fewer than 10 included studies.
✔ Yes
☐ No
☐ Can't answer
☐ Not applicable

11. Was the conflict of interest included?
Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.
Note: To get a “yes,” must indicate source of funding or support for the systematic review AND for each of the included studies.
☐ Yes
☐ No
☐ Can't answer
✔ Not applicable

Additional notes (in italics) made by Michelle Weir, Julia Worswick, and Carolyn Wayne based on conversations with
Bev Shea and/or Jeremy Grimshaw in June and October 2008 and July and September 2010
Do impairments in facial affect recognition ability lead to psychotic symptoms?
A systematic review and meta-analysis

Natalie Bordon\textsuperscript{a1}, Suzanne O’Rourke\textsuperscript{2} and Paul Hutton\textsuperscript{2}

\textsuperscript{a} Corresponding author email address; Natalie.Bordon@nhs.net
\textsuperscript{1} The State Hospital, Lanark, NHS The State Hospital.
\textsuperscript{2} The School of Health in Social Science, the University of Edinburgh

Prepared for submission to Psychological Review
Abstract
There is a substantial evidence base demonstrating difficulties in emotional perception in patients with psychosis, with evidence suggesting a relationship with reduced social functioning, increased aggression and more severe symptoms of psychosis. In this review we aim to review this field to assess if there is a causal link between facial affect recognition difficulties and psychosis. The Bradford Hill criteria for establishing a causal relationship from observational data were used to generate key hypotheses, which were then tested against existing evidence. Where a published meta-analysis was not already available, new meta-analyses were conducted. A large effect of FAR difficulties in those with a diagnosis of psychosis, with a small to moderate correlation between FAR problems and symptoms of psychosis was found. Evidence was provided for the existence of FAR problems in those at clinical high risk of psychosis, while remediation of psychosis symptoms did not appear to impact FAR difficulties. There appears to be good evidence of the existence of facial affect recognition difficulties in the causation of psychosis, though larger, longitudinal studies are required to provide further evidence of this.
Introduction
There is good evidence for an association between having a diagnosis of psychosis and difficulties in social functioning (Cramer, Bowen, & O'Neill, 1992; Kerr & Neale, 1993; Sullivan & Allen, 1999). Some authors argue that this may be mediated by poor social cognitive processing in psychosis, more specifically facial affect recognition (FAR) (Blair, 1995; Malone, Carroll, & Murphy, 2012). There is evidence that patients with a diagnosis of psychosis have difficulties with facial affect recognition (Edwards, Jackson, & Pattison, 2002; Kohler, Walker, Martin, Healey, & Moberg, 2009; Mandal, Pandey, & Prasad, 1998; R. W. Morris, Weickert, & Loughland, 2009) and that these difficulties are associated with reduced social functioning (Hooker & Park, 2002; Irani, Seligman, Kamath, Kohler, & Gur, 2012) and increased aggression (Frommann, Stroth, Brinkmeyer, Wolwer, & Luckhaus, 2013; Hoaken, Allaby, & Earle, 2007). FAR impairments may also contribute to increased psychotic symptoms, such as paranoia, asociality and anhedonia. Difficulty in interpreting emotions correctly could generate confusion regarding the intentions of others, which may heighten fears of being harmed (M. J. Green & Phillips, 2004; Kohler et al., 2003), social avoidance (Eack et al., 2009) and lead to difficulty in forming secure attachments (Steele, Steele, & Croft, 2008; Suslow et al., 2009).

Although it is theoretically plausible that FAR difficulties could directly contribute to psychotic symptoms, the empirical evidence is far from clear. Whether problems with FAR are a cause, consequence or non-specific correlate of these experiences is unknown. The aim of this paper is therefore to systematically review the extensive evidence on FAR and psychosis, and then test it against a set of hypotheses derived from the Bradford Hill Criteria (Hill, 1965) for establishing causality. This will not only clarify whether and to what extent credible claims can, at present, be made for FAR impairments being a possible pathway to psychotic symptoms, it will also identify the key gaps in the literature that need to be filled before firmer conclusions can be drawn. Before outlining the specific hypotheses we intend to examine, we will first consider in more detail existing cognitive models of psychosis, and the implications of these for considering how FAR impairments and psychotic symptoms might be related.

Facial affect recognition and cognitive models of psychosis
Cognitive models of the positive symptoms of psychosis (Garety, Kuipers, Fowler, Freeman, & Bebbington, 2001), highlight that psychosis occurs in people with a vulnerable predisposition and that onset can be triggered by life events. Emotional changes can occur in addition to disruptions or biases in cognitive processes such as attention, perception or judgment. Several biased appraisal processes have been identified including jumping to conclusions (Garety, Hemsley, & Wessely, 1991), external attributional biases (Richard P. Bentall, Corcoran, Howard, Blackwood, & Kinderman, 2001; Kinderman & Bentall, 1997) and difficulties in understanding social situations (Bora & Pantelis, 2013; Brüne, 2005).

In terms of the development of the positive symptoms of psychosis, two routes have been proposed, one in which psychosis develops through cognitive and affect changes, and one in which it develops through affect disturbances alone (Garety et al., 2001). Garety and
colleagues (2001) proposed that anomalous conscious experiences such as heightened perception or racing thoughts are triggered by these cognitive disturbances. The affected individual, in response to finding them strange and uncomfortable, endeavours to explain their presence – this ‘search for meaning’ may often, perhaps because of pre-existing cognitive biases, conclude with the individual attributing the cause of these experiences to an external source. The resulting explanation may, if it meets certain criteria, be classified as a delusion, conviction in which may be maintained by a variety of cognitive, behavioural and interpersonal processes, each of which serve to prevent disconfirmation (Garety et al., 2001).

The person is also likely to experience emotional changes such as increased anxiety, due to both the triggering event and also their reaction to the anomalous experiences. These emotional changes can affect the processing of the anomalous experiences, which due to their nature are puzzling for the individual, and therefore also trigger a search for an explanation (Maher, 1988).

The role of social cognition has also been investigated in psychosis. This refers to the way in which we understand, perceive and interpret our social world (David L Penn, Corrigan, Bentall, Racenstein, & Newman, 1997) and is made up of various constructs including facial affect recognition, ‘theory of mind’, social perception and our ability to make appropriate attributions for events. It has been suggested that facial affect recognition impairments may lead to a confusing social world for people with psychosis, which may increase anomalous experiences (M. J. Green & Phillips, 2004; Poole, Tobias, & Vinogradov, 2000). Attempting to make sense of this could trigger an increase in positive symptoms such as paranoia (Garety et al., 2001; M. J. Green & Phillips, 2004) and delusional ideation (Arguedas, Green, Langdon, & Coltheart, 2006). As Garety et al (2001) suggest, ‘biased conscious experiences’ are important in so far contribute to the belief that the anomalous experiences are external in nature. The inability to correctly perceive the emotional states of others from their facial expression may lead to confusion regarding the intentions of others which, in combination with reasoning biases and other factors, could contribute to causing, exacerbating or maintaining delusional ideation.

Early adversity, facial affect recognition and positive symptoms of psychosis

Psychological and biological researchers are in general agreement that early trauma can considerably increase the risk of psychosis in adult life (Barker, Gumley, Schwannauer, & Lawrie, 2015). Varese et al. (2012) found significant associations between childhood adversities and psychosis (OR=2.78) and reported that those with psychosis were 2.72 times more likely to have been exposed to childhood adversity than controls. Bentall and colleagues (2012) propose that different types of trauma are likely to have different effects on different cognitive processes, which may explain why different types of trauma and adversity are more associated with certain symptoms of psychosis than others. For example, childhood rape is associated with hallucinations when controlling for co-occurring paranoia (OR=8.9), whereas being placed in care is associated with paranoia when controlling for hallucinations (OR=11.08) (Richard P. Bentall, Wickham, Shevlin, & Varese, 2012). This raises the possibility that some types of early adversity are associated with psychosis via their effects on
facial affect recognition ability. We will now consider whether neglect, disrupted attachment and paranoia may be involved in one such pathway.

The ability to recognise FAR does not emerge at one particular stage of development but appears to develop gradually over time. The ability to recognise happiness appears to develop first, followed by sadness or anger, and then fear and surprise (De Sonneville et al., 2002; Gross & Ballif, 1991; Smith & Walden, 1998; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000). Parents presumably have a crucial role in the development of this ability. By repeatedly labelling displays of affect and providing exposure to different types of emotion, they can strengthen FAR ability throughout development (Beale & Keil, 1995; Keyes, 2012; Pollak, 2003; Shenk, Putnam, & Noll, 2013). In the absence of this process, children have been shown to be less emotionally competent (Denham, Mitchell-Copeland, Strandberg, Auerbach, & Blair, 1997; Fruzzetti & Shenk, 2008) and Fries et al (2005) found that children who have experienced disrupted attachment also have a specifically impaired ability to distinguish between emotions (Fries, Ziegler, Kurian, Jacoris, & Pollak, 2005). Given the findings of R. P. Bentall, Rowse, Shryane, and et al. (2009), that disruption of early attachments and childhood adversity may be associated with an over-anticipation of social threats, one interesting hypothesis may be that impaired FAR ability may mediate the effect of disrupted attachment on paranoia.

There is another way in which trauma may affect emotion recognition and subsequently increase the risk of psychotic experiences. Pollak (2008) suggested that traumatic early experiences alter sensory thresholds for particular emotions. For example, children who have been physically abused demonstrated a heightened ability to detect angry faces and were more likely to label ambiguous faces as angry than non-abused children (Pollak & Kistler, 2002). Smith and Walden (1998) reported that children from more deprived socio-economic backgrounds were more accurate in the recognition of the expression of fear than those from more advantaged socio-economic households. The authors propose that when growing up in a high-stress environment recognising the expression of fear would be important to help avoid potentially dangerous situations, and may also be a more common emotion to not only witness but also experience - meaning these children have greater experience and knowledge of this emotion. It may be that a heightened sensitivity to expressions of fear and anger, as well as reduced exposure to a range of more positive emotions, may predispose such individuals to more distressing fears or being harmed. Contradicting this, however, are the results of studies suggesting that people with psychosis have greater difficulty recognising negative facial expressions, compared to positive ones (Demirbuga et al., 2013; Hofer et al., 2009).

Although the evidence above suggests some plausible mechanisms linking FAR to paranoia and possibly other positive symptoms of psychosis (eg., delusions of reference), more research is needed to test these emerging hypotheses. However the contribution of FAR impairments to impaired social functioning and other so-called negative symptoms that characterise psychosis has also been the subject of considerable research, which is important given the lack of effective psychosocial or pharmacological treatments for these experiences.
Early adversity, facial affect recognition and negative symptoms of psychosis

It is well-documented that patients with psychosis have impaired social functioning (APA, 1994; Bellack, Morrison, Wixted, & Mueser, 1990). Although social exclusion, stigma and social anxiety no doubt play a large part in this, impaired social cognition may also contribute to this (Jean Addington, Saeedi, & Addington, 2006b; David L Penn et al., 1997), as it does in other disorders. For example, emotion recognition and mental state recognition are known to be impaired in people with a diagnosis of autistic spectrum disorder (ASD), and it is generally accepted that this contributes a great deal to the social difficulties they experience (Baron-Cohen, 1997; Hobson, 2013). Accurate recognition of facial affect is a highly complex task, and impairments in this area are likely to have a direct and substantial impact on our ability to interact socially (Couture, Penn, & Roberts, 2006) and function in the world (Irani et al., 2012; Pan, Chen, Chen, & Liu, 2009). In psychosis, impaired facial affect recognition has been associated with asociality (Poole et al., 2000), impaired emotional expression (Gaebel & Wölwer, 1992), anhedonia (M. Green & Walker, 1986; Gur et al., 2006; Neale, Oltmanns, & Harvey, 1985), poor performance in social role plays and a choice of personal appearance that is judged by others in one’s culture to be socially inappropriate (Allott et al., 2014; Mueser et al., 1996).

Is FAR impairment one cause of psychosis?

Although FAR ability has been associated with a variety of symptoms and risk factors for psychosis, and although it is plausible that FAR plays a causal or maintaining role in these, the empirical evidence for such a claim has yet to be systematically assessed. One widely used set of criteria for determining the strength of evidence for such claims is the Bradford Hill Criteria (Hill, 1965). These criteria were initially developed as a way of determining a causal link between a specific factor and a disease, namely smoking and lung cancer, and have been widely used as a way of establishing scientifically valid casual connections in epidemiological research. The criteria have also been applied to assess mechanisms of change in psychotherapy research (Kazdin, 2007) and provide a set of minimal conditions necessary for evidencing a causal relationship between two factors.

We will now consider each of the criteria in turn, and how these have directed the hypotheses against which we will test the evidence.

Strength

According to Bradford Hill, causal relationships are more likely to produce strong relationships than weak ones. Our first hypothesis was therefore that FAR impairments will be strongly correlated with psychotic and negative symptoms.

Consistency

Causal relationships are more likely than non-causal ones to produce findings that are not moderated by differences in contexts, populations and time periods. Our second hypothesis was therefore that there would be low heterogeneity in meta-analyses examining the relationship between FAR and psychosis, as assessed by the $I^2$ statistic.
Specificity
Bradford Hill’s third observation was that causality is more likely when the hypothesised causal variable is only, or specifically, associated with the hypothesised cause. Not all causal variables have specific effects, but if such a pattern is observed, it increases the probability that a causal relationship exists. We therefore sought to test the hypotheses that FAR ability would be more impaired in people with psychosis than people with non-psychotic mental health problems, and that people with FAR impairments would be more likely to develop psychosis.

Temporality
The fourth criterion is temporality as in order to infer a causal relationship, the cause must precede the effect. If FAR difficulties have a causal role in psychosis, to establish temporality, FAR problems must be present prior to onset of psychotic symptoms.

Biological gradient (dose-response)
Biological gradient refers to the concept that demonstration of a gradient in which greater levels of a cause is associated with greater levels in the outcome. Evidence of a ‘dose-response’ relationship may be provided by those individuals who have experience FAR difficulties for a longer period of time also having more severe psychotic symptoms. Additionally if FAR difficulties are modified across low, moderate and high intensities, similar levels of remission would also be seen in psychotic symptoms if a dose-response relationship was present.

Plausibility and coherence
Bradford Hill suggests that the plausibility and coherence of how well an explanation for a mechanism fits within broader evidence and theories should be examined, therefore FAR difficulties within the psychosis population will be explored in relation to current theories of psychosis.

Experimental evidence
An experiment criterion in which direct manipulation of a proposed causal factor impacts the outcome also contributes evidence to a causal relationship between the two. For the purposes of this review, assessment of the modification of FAR abilities on symptoms of psychosis will be explored.

Analogy (compelling alternative explanations)
Bradford Hill proposes that as part of an assessment of causality, it is important to understand the extent to which other possible explanations have been considered and ruled out. A more general face-processing difficulty or problems with attention may explain the association between FAR problems and psychosis and will be discussed.

Method
Operationalisation of the criteria
For each of the hypotheses, if there was a meta-analysis published within the last 6 years containing more than 15 studies with a minimum of 700 participants, then this evidence was used. If not, an up to date meta-analysis was conducted where sufficient studies were available. Research recommendations were made in the absence of sufficient evidence.

**Search Strategy**
A systematic review of the literature was conducted in accordance with PRISMA and AMSTAR guidelines. A search of the following electronic databases was carried out in March 2016: Medline, Embase, PsychInfo and Web Science. All years available were searched, using the following terms: ‘facial affect recognition, facial emotion recognition, facial affect recognition training, social cognition, emotion perception, schizophrenia’ and ‘schizoaffective’. Additionally in order to identify any unpublished studies the US government clinical trials register (clinicaltrials.gov), European Union clinical trials register (clinicaltrials-register.eu), World Health Organisation (apps.who.int/trialsearch) and Current Controlled Trials Ltd (controlled-trials.com) were all searched in March 2016. Reference sections within the articles which met the inclusion criteria were also searched by hand to identify any further papers.

**Inclusion Criteria**
Eligible studies required at least 50% of participants in at least one of the study groups to have a diagnosis of non-affective psychosis. Studies where 50% or more participants in the psychosis group also had learning disability, predominantly substance induced psychosis, or organic brain damage were excluded. Studies which took place in a variety of settings such as inpatient and outpatient were included providing the other criteria were met.

For intervention studies, eligible studies had to assess the effect of interventions that were specifically designed to improve facial affect recognition. Studies were only included if more than 50% of the intervention was specifically targeting facial affect recognition. This was determined by accessing the intervention manual or description, and calculating a percentage of the total time dedicated within the programme to facial affect recognition training. In order to minimise risk of bias, only randomised controlled studies were included in the meta-analyses.

**Data Extraction and Outcomes**
In line with previous meta-analyses (Kohler et al., 2009; Kurtz & Richardson, 2012) different measures of facial affect identification and discrimination were included and combined given the assumed similarity of the task of labelling emotional expressions and distinguishing emotional expressions between two different faces. Some of the most common tasks included the Facial Emotion Identification Test (FEIT), Penn Emotion Recognition Task (ER40), Pictures of Facial Affect (PFA) and Bell Lysaker Emotion Recogntion Scale (BLERT). Many of these measures are based on the Ekman pictures of facial affect (Ekman & Friesen, 1976) making their results comparable. Only studies using these or other valid and reliable measures of facial affect recognition were eligible for inclusion. The BLERT and ER40 have previously been shown to have good psychometric properties for this patient group (Pinkham,
Penn, Green, & Harvey, 2015) though it is acknowledged that further assessment of social cognitive measures for the psychosis population is required. For symptoms, we used data derived from the PANSS or the BPRS or any other reliable and valid measure of symptoms as used by study authors. Authors were contacted in the event of missing data.

Meta-Analysis Calculations
If studies had two or more similar arms, then these were combined into one using procedures outlined in the Cochrane Handbook (Higgins et al., 2011). For each meta-analysis, means and associated standard deviations were entered into MetaXL (Barendregt & Doi) or reported statistics were transformed into effect sizes to allow a computed pooled standardised mean difference (Hedges’s g) and 95% confidence interval to be calculated. Given that the observational studies and RCTs included in the analysis used different measures and there may be baseline differences, standardised mean difference was chosen as this allows comparison across studies. Where repeated measure designs were used, effect sizes were calculated in line with procedures outlined in Peters and Mengersen (2008), S. B. Morris (2000), S. B. Morris and DeShon (2002) and Earl and Albarracín (2007). A random effects model was used, since this assumes the true effect size can vary across studies (Borenstein, Hedges, Higgins, & Rothstein, 2009), and that the individual effect sizes are a random sample from the distribution of possible effects. The heterogeneity of the effect sizes was measured using the $I^2$ statistic, and a Chi-Square test was performed to evaluate if the intervention effects vary more than could be expected due to random error only.

Study Quality
The AHRQ (Agency for Healthcare Research and Quality) (Health & Services, 2012) was used to assess the quality of the studies. This tool is recommended by the CRD (Centre for Reviews and Dissemination)(2009) as suitable for assessing the quality of observational studies and it is advised that the tool should be adapted for the individual requirements of the systematic review. An adapted version of the AHRQ was therefore used (Appendix 2) which included the domains of selection bias, detection bias, statistical power, validity of measures and method of analysis. Each item within the domains was rated using the tool and assigned a rating of either ‘yes’, ‘no’, ‘partially’, ‘or unclear’.

For the RCT studies, The Cochrane Collaboration Risk of Bias Tool (Higgins et al., 2011) was used to assess the methodological quality of the RCT trials included in the meta-analysis. It involves examining a range of biases that can occur in trials such as how participants are randomised, blinding of both participants and study personnel and selective reporting of results. Each feature of interest is given a rating of either ‘low’, ‘high’ or ‘unclear’ risk of bias and these ratings are then taken into account when interpreting the effect sizes of the outcomes and subsequently the conclusions that can be made from the data.

Results
Study Selection
Figures 1 to 4 outline the process of study selection for each meta-analysis carried out. The initial search, after removal of duplicates, identified 2835 papers, conference abstracts and
dissertations. The majority of these papers were discarded on the basis of their title where it was clear that they did not include people with psychosis, specific measures of FAR or FRT. Abstracts of the remaining papers were then reviewed allowing further papers to be discarded on the basis that they did not meet inclusion criteria from the information contained in the abstract. Finally the full texts of the remaining papers were reviewed and papers were identified as suitable for inclusion within each of the meta-analyses. The reference sections of these papers were then hand searched to identify any further papers suitable for inclusion. Where papers were found using this method, this is outlined in the search diagram (Figures 1 to 4).

Study Characteristics
A total of 10 meta-analyses on aspects of facial affect recognition in the psychosis population were identified and 2 met the pre-specified inclusion criteria. For a comparison of non-affective and affective psychosis, a total of 10 studies involving 667 participants were included in the analysis as shown in Table 1. Nine studies (1649 participants) were included in the meta-analysis of FAR in people with active and non-active psychosis (Table 2) and 4 studies (933 participants) were included in the meta-analysis investigating FAR in the clinical high risk population (Table 3). All studies used a variety of commonly used measures such as the PANSS, SAPS and SANS, to assess psychotic symptoms.

A total of 8 studies involving 300 participants were included in the meta-analysis investigating FRT programmes. As shown in Table 4, a range of intervention programmes were assessed. These included Training of Affect Recognition, Attentional Shaping, Micro-Expression Training Tool and Facial Feedback. All programmes were solely focused on improving facial affect recognition, but varied in duration from 1 treatment session to 12 sessions. Control group participants received various interventions, including cognitive remediation therapy, repeated exposure to pictures of facial affect, or simply treatment as usual.

The evidence will now be tested against each of the Bradford Hill derived-hypotheses:

Strength
A causal effect of FAR problems on psychosis would be consistent with there being a strong association between FAR and psychotic symptoms. We found two recent meta-analyses that have already examined this question (Kohler et al., 2009; J. Ventura, Wood, & Hellemann, 2013). Kohler et al. (2009) found that participants with psychosis were much less able to accurately identify facial affect than non-clinical individuals without psychosis (N=3822, k=86, d=-0.91, 95% CI -0.97, -0.84, P= 74%). Joseph Ventura, Wood, Jimenez, and Hellemann (2013) reported small to moderate negative correlations between FAR ability and negative symptoms (N=2303, k=53, r= -0.25, P= 55%), positive symptoms (N=771, k=17, r= -0.17, P= 36%), reality distortion (N=757, k=18, r= -0.21, P=60%) and conceptual disorganisation (N=987, k=22, r= -0.32, P=44%).

Consistency
While Kohler et al found a large precise effect between FAR difficulties and having a diagnosis of psychosis there was considerable heterogeneity within the sample ($I^2$ 74%; Kohler et al., 2009). Inspection of forest plots, however, suggests the heterogeneity related to the magnitude of the difference rather than its presence. Only 15 of the 86 included studies did not find a significant reduction in FAR ability in people with psychosis (Kohler et al., 2009). While small to moderate correlations were established between psychotic symptoms and FAR difficulties, these findings were more consistent with moderate heterogeneity in the samples (Venture et al, 2013). The most consistent evidence was for positive symptoms ($I^2$ 33%; Ventura et al, 2013).

Inconsistency was also assessed by inspection of the degree of heterogeneity in our meta-analytical findings (see Table 5). The estimates of the improvement in FAR ability immediately after FRT and at one-week follow up, showed substantial heterogeneity ($P$ 65% and 72% respectively), as did the estimate of the effect of FRT when compared to active control groups ($P$ 76%). The comparison of FAR in non-affective and affective psychosis suffered from very high heterogeneity ($P$ 98%). Analyses of the impact of active symptoms of psychosis on FAR ability and also estimates of FAR in the clinical high risk population, were both free of heterogeneity.

**Specificity**

In order to examine specificity of the FAR impairment to non-affective psychosis, a meta-analysis of 10 studies (667 participants) comparing FAR ability in people with a diagnosis of non-affective psychosis against people with a diagnosis of affective psychosis/bipolar disorder was carried out. This found a moderate effect size ($N$=667, $k$=10, $g$=-0.62, 95% CI -1.21, -0.03, $P$=98%; see Figure 5), suggesting patients with non-affective psychosis have significantly lower FAR ability than those with a diagnosis of affective psychosis. There was substantial heterogeneity within the sample, which may be attributed to the very large effect sizes reported by Lee et al. (2013) and Thaler, Allen, Sutton, Vertinski, and Ringdahl (2013). Furthermore given the poor quality of some of the included studies and imprecision of the effect size, this evidence was judged to be of very low quality.

**Temporality**

If reduced FAR ability causes psychotic symptoms, then this suggests it should be present before psychotic symptoms develop. If this is the case, then FAR difficulties of a magnitude seen in established psychosis should be evident in individuals who have not developed psychosis but are at risk of doing so – i.e., those who meet clinical high risk criteria. As shown in Figures 6(a) and 6(b), we conducted two meta-analyses, one of which found that participants at high risk of developing psychosis had significantly poorer FAR ability compared to healthy individuals ($N$=1291, $k$=3, $g$=-0.35, 95% CI -0.38, -0.32, $P$=0%) and another found no difference in FAR ability between those CHR individuals who later converted to psychosis compared to those who did not ($N$=933, $k$=4, $g$=-0.07, 95% CI -0.19, 0.06, $P$=0%; see Figure 6(b)). If FAR problems cause psychosis, then it should also be the case that currently healthy people with FAR problems are more likely to develop psychosis in
the future, than those without them. Unfortunately no studies reported data relevant to this question.

Finally, if FAR difficulties resolve after psychotic symptoms are successfully treated, then this would support the competing hypothesis that FAR difficulties are a consequence or artefact of psychotic symptoms. We therefore carried out a meta-analysis of 9 longitudinal studies (6 repeated measures; 3 independent groups) as shown in Figure 7, and found no evidence that successful treatment of psychotic symptoms was associated with resolution of FAR difficulties (N=780, k=9, g=0.02, 95% CI -0.18, 0.22, P=0%).

**Biological Gradient**

There are at least two ways to examine whether greater exposure to FAR difficulties leads to greater severity of psychotic symptoms. First, if there is such a ‘dose-response’ relationship, then those who have been exposed to untreated FAR difficulties for longer should have more severe psychotic symptoms. Second, those randomised to FAR-enhancing interventions of low, moderate and high intensities or durations ought to experience comparable reductions in psychotic symptoms. There were no relevant studies. Although several trials of FAR-enhancing interventions have been carried out (see below), these have not examined the effect of varying the intensity of the intervention on psychotic symptoms.

**Plausibility and Coherence**

Given the evidence previously discussed in relation to the childhood adversity model of psychosis (Varese et al., 2012) and also the impact of child maltreatment on FAR (Richard P. Bentall et al., 2012), it is proposed that there may be common aetiology between the development of both psychosis and FAR difficulties. Being brought up within an environment where care givers are not responsive to the child’s needs, or the child is neglected has established links to later development of psychosis (Read, Fink, Rudegeair, Felitti, & Whitfield, 2008; Read, Os, Morrison, & Ross, 2005; Whitfield, Dube, Felitti, & Anda, 2005) and could also result in the child having difficulty recognising and understanding the emotions of others due to lack of exposure to these emotions in others (Beale & Keil, 1995; Keyes, 2012; Pollak, 2003). Children who have been abused are more at risk of developing psychosis (Varese et al., 2012) and are more likely to over-anticipate social threat from others (Richard P. Bentall et al., 2012). It is plausible that this relationship could be mediated by FAR abilities; that is, neglect and abuse could distort the development of FAR ability, leading to misidentification of emotions and the intentions of others.

Furthermore it is proposed that the FAR problems would fit within current cognitive theories of psychosis (Freeman, Garety, Kuipers, Fowler, & Bebbington, 2002; Garety et al., 2001). In their review of delusions, Freeman and Garety (2014) highlight that patients with current symptoms of psychosis have difficulties with theory of mind. Given that the first stage of this process could be thought of as correctly recognising the facial expression of another person, before it can be processed to infer the emotional state of them, FAR would fit within this established reasoning bias (Brüne, 2005; Frith, 1994; Sprong, Schotborst, Vos, Hox, & Van Engeland, 2007). Additionally FAR problems could prevent individuals from interacting
competently in a social world as if one cannot read emotions in faces, then this could cause considerable confusion and uncertainty. This odd experience could trigger a search for meaning and if someone also has a reasoning bias – eg they form conclusions quickly – then the conclusion they form might end up being a delusional one. FAR could therefore also mediate the reasoning biases that a person may have due to their psychosis.

Experimental evidence
Observational data is subject to numerous confounds (Austin, 2011; Mann & Wood, 2012; Viswanathan, Berkman, Dryden, & Hartling, 2013) and, as Bradford Hill (1965) notes, experimental evidence can often provide the strongest support for a causation hypothesis. It is possible to experimentally (and ethically) test the effect of FAR on psychotic symptoms by randomly allocating participants to a FAR-enhancing intervention or a control condition. It may also be possible to randomly allocate healthy individuals to a mild FAR-inducing intervention or a control intervention and see whether the induction causes a mild increase in subclinical psychotic symptoms.

In order to assess the impact of reduction of the FAR impairment on psychosis symptoms, first of all a meta-analysis was conducted to assess whether this could be significantly modified and, should a reduction in FAR ability be found, its impact on psychosis symptoms explored. An analysis of post-intervention data from 8 RCTs found a very large effect of FRT on facial affect recognition ability \[N=300, k=8, g=1.35, 95\% \ CI 0.89, 1.80, P=65\%; \text{see Figure 8 (a)}\]. Three studies (108 participants) reported data at 1-week following the intervention. As shown in Figure 8 (b), a large significant effect was found \[N=108, k=3, g=1.46, 95\% \ CI 0.61, 2.32, P=72\%\], however there was considerable heterogeneity. Only one study (Luckhaus, Frommann, Stroth, Brinkmeyer, & Wolwer, 2013) provided 8-week follow-up data, and found no change in affect recognition scores at follow-up from those taken immediately post intervention \[T=1.210, df=14; p=0.246\] may suggest that the treatment effect was maintained, however given the lack of statistical power this result should be interpreted with some caution.

A third analysis was carried out of studies comparing FRT to interventions involving repeated exposure to pictures of facial affect or Cognitive Remediation Training (CRT), since these control for non-specific effects of additional therapeutic attention and time [see Figure 8(c)]. Comparisons involving waiting list or usual treatment groups were excluded from this analysis. This calculation was based on 5 studies and included data for 198 participants. As with the previous analyses, a large significant effect was found \[N=198, k=5, g=1.60, 95\% \ CI 0.91, 2.28, P=72\%\] although substantial heterogeneity was again observed.

The above analyses suggest FRT causes reliable and large improvements in FAR ability in psychosis, and that these effects are not an artefact of non-specific therapeutic factors. Thus, if FRT also causes improvement in psychotic symptoms, it is reasonable to assume this is likely to be attributable to improved FAR ability. If results of randomised controlled trials show that intervening to improve FAR ability causes improvements in psychotic symptoms, then this would be strong evidence for FAR having a causal relationship with psychotic
symptoms. To test this hypothesis, three meta-analyses of randomised controlled trials were conducted. These found no significant effects of FRT on negative symptoms \([N=173, k=4; g=-0.11, 95\% \text{ CI} -0.41, 0.20, P=0\%; \text{see Figure } 9(a)]\), positive symptoms \([N=135, k=3; g=0.10, 95\% \text{ CI} -0.25, 0.45, P=0\%; \text{see Figure } 9(b)]\), or general psychopathology \([N=135, k=3; g=0.12, 95\% \text{ CI} -0.44, 0.69, P=56\%; \text{see Figure } 9(c)]\). Only two uncontrolled studies measured and reported data on the effect of FRT on psychosis symptoms (Drusch, Stroth, Kamp, Frommann, & Wolwer, 2014; Frommann, Streit, & Wolwer, 2003). These studies reported conflicting results, with Frommann et al. (2003) reporting a significant improvement in symptoms following FRT, and Drusch et al. (2014) finding no change.

In regards to the quality of the RCT studies included in the analyses, the most significant limitation was in regards to the randomisation process and blinding of personnel and participants. No information was contained in the articles regarding how participants were assigned to the intervention and control groups and the majority of the studies did not report that assessors were blinded to which group participants were part of. Although lead authors were contacted and asked for further information regarding this, in the majority of cases no further information was provided.

**Analogy (alternative explanations)**

Previous studies have demonstrated that patients with a diagnosis of psychosis appear to focus on the non-salient parts of the face that would help identify affect (Phillips & David, 1997; Streit, Wolwer, & Gaebel, 1997). It has also been established that patients with a diagnosis of psychosis perform slightly worse than healthy controls in processing non-emotional facial features like identity (Archer, Hay, & Young, 1992; Bediou et al., 2005; Kerr & Neale, 1993), therefore perhaps a more general difficulty in face processing is responsible for the association between FAR and psychosis. Visual attentional impairments have been reported in patients with psychosis (Brenner, Lysaker, Wilt, & O'Donnell, 2002; Butler & Javitt, 2005; Butler et al., 2001) therefore this may also mediate FAR ability. However Joseph Ventura et al. (2013) reported that while neurocognition and social cognition were related constructs, no one specific neurocognitive function was specifically associated with emotional processing in order to account for the observed impairments.

**Discussion**

**Summary of Findings**

Kohler et al (2007) found that FAR difficulties have a strong association with psychosis. However the large precise effect size these authors reported, must be taken in the context of the small to moderate associations between FAR problems and symptoms of psychosis (Ventura et al 2013). Furthermore this evidence is affected by the reduced quality of the included studies, mainly due to small sample sizes and subsequent lack of power in these studies, and also moderate heterogeneity. There was some variation in the consistency across outcomes with heterogeneity ranging from absent to substantial, although most fell within the
moderate range. In terms of specificity, there appeared to be some evidence of FAR problems being more prevalent in those with non-affective psychosis than those with affective psychosis, however comparisons of FAR ability in those with a diagnosis of psychosis with other mental health problems would add further evidence to support this hypothesis. Once again this outcome is deemed low quality due to small sample sizes, lack of coherent reporting of results in the included papers and substantial heterogeneity.

To establish temporality there appeared to be support for those at clinical high risk of developing psychosis displaying FAR difficulties when compared to healthy controls, however this was not of the magnitude reported in those with a diagnosis of psychosis and there was no significant differences between those who later converted to psychosis and those who did not. However given that FAR problems do not resolve when psychosis symptoms are treated, then support for the competing hypothesis can also not be supported. Additionally no evidence could be found to test the biological gradient criteria and given the few studies that reported the impact of FRT programmes on psychotic symptoms, presently further studies would be required to definitively answer the experimental hypothesis. In relation to problems with FAR abilities being associated with psychosis, integration into current theories and models appears both plausible and coherent. While an alternative explanation of the association between FAR ability and psychosis being caused by attentional deficits may be proposed, there appeared to be some evidence that attentional difficulties were not associated with social cognitive processes.

**Does FAR cause psychosis?**

On reviewing each of the outlined Bradford Hill Criteria (Hill, 1965) there appears to be strong evidence for the existence of FAR difficulties within the psychosis population and furthermore that these problems are associated with psychotic symptoms. The significant correlation between FAR difficulties and symptoms of psychosis, together with evidence that these problems precede the onset of psychosis (ie., are present in those assessed as being at clinical high risk of psychosis), suggests that FAR may be an underlying vulnerability that predisposes people to psychotic symptoms – i.e., that FAR impairments may be one cause of psychotic symptoms. On the other hand, it could be argued that FAR impairments are a consequence of psychotic symptoms - however this theory would struggle to explain why FAR impairments precede the onset of psychosis and continue when they subside.

Interestingly, FAR difficulties appear to be moderately more pronounced in people with non-affective psychosis than those with affective psychosis, which is consistent with a degree of specificity. However no studies provided matched data from people with other mental health problems, meaning it remains possible that FAR difficulties are related to mental illness generally – rather than psychosis specifically. Although there is a correlation between FAR ability and psychotic symptoms, the observation that interventions that enhance FAR do not seem to ameliorate psychotic symptoms does challenge the claim that FAR has a specific causal role in psychosis. On the other hand, most of the trials were very brief, and it may be that FAR-enhancing interventions need more time to have a measurable antipsychotic effect.
A range of reasoning and attributional biases are thought to be associated with psychosis. Might the difficulties this group seem to have in recognising emotion also be conceptualised as a bias? What constitutes a ‘bias’ rather than an ‘impairment’ is unclear but regardless, the degree to which people with psychosis struggle to recognise facial affect is remarkable. For context, consider the well-known ‘jumping to conclusions’ (JTC) bias - a reasoning bias which features highly in cognitive accounts of psychosis. According to the most recent meta-analysis, people without psychosis request moderately more information before making a decision than people with psychosis, (N=1935, k=33, g= -0.53, 95% CI -0.69,-0.36, P=66%) (Dudley, Taylor, Wickham, & Hutton, 2015). With respect to FAR ability, however, people with psychosis have a very large disadvantage compared to those without psychosis (N=3822, k=86, d= -0.91, 95% CI - 0.97, -0.84, P= 74%) (Kohler et al., 2009). It is of course possible that these two phenomena interact – perhaps those with both a JTC bias and difficulty recognising emotions are more vulnerable to psychosis than those with just one of these features? Future studies testing for such interactions would be welcome.

Research recommendations
To help establish temporality, longitudinal studies are required that identify those who have FAR difficulties and follow them up over time to investigate if they are more likely to develop psychosis when compared to those who do not have FAR difficulties. While it seems reasonable to conclude that modification of facial affect recognition ability is possible in psychosis given the large improvements in FAR, existing studies are small and there is a need for large, rater-blind trials. In order to provide evidence for a dose-response relationship, future randomised trials should be conducted that provide FAR-enhancing interventions to varying intensities to investigate if this coincides with different levels of improvement in psychotic symptoms.

A wider recommendation for future research would be the measurement and reporting of outcomes that may be considered more clinically relevant and related to a participants’ functioning in the real world, such as symptoms of psychosis or social functioning. While an improvement in a patient’s ability to recognise emotional expressions is positive, it is perhaps more important to investigate if this indeed benefits them in their daily interactions with others and subsequently improves their functioning or quality of life. Additionally as there are theories to suggest that misinterpretation of facial expressions in others (Kohler et al., 2003; Lemerise & Arsenio, 2000) can lead to aggression and violence, the effectiveness of these programmes in relation to this outcome would be of significant relevance to those detained within forensic mental health settings to explore if it reduces their risk of future violence.

Limitations
While there appears to be good evidence for an existence of a FAR difficulties in those with a diagnosis of psychosis and a high quality review of the evidence was conducted, some caution must be taken when interpreting these results. The AMSTAR (Shea et al., 2007) quality assessment criteria was used to measure the quality of this systematic review. This
tool consists of 11 concise criterion items and a score of 1 is given if the criterion is met and a score of 0 if it is not met, unclear or not applicable. A score of 8 resulted indicating a high quality review (Sharif, Janjua-Sharif, Ali, & Ahmed, 2013). In accordance with the GRADE approach (2004), the quality of the outcomes were assessed as low quality for the outcomes of association of FAR problems and psychosis, improvement in FAR after FRT, and FAR ability in those at CHR for psychosis who convert to psychosis and very low quality for the remaining outcomes. The main area that reduced the quality of the evidence for the outcome measures was the quality of the included studies mainly due to small sample sizes, reduced power and inconsistent reporting of results. Furthermore not all studies reported data that was compatible with conducting a meta-analysis. In these cases authors of the original papers were contacted, however not all responded therefore some studies were excluded on this basis.

**Conclusion and implications**

The importance of the role social cognition plays within the recovery model of psychosis has been highlighted previous (Gumley, Braehler, & Macbeth, 2014) and recommendations made that relapse prevention efforts focus on ‘social relating, their developmental origins and interpersonal context rather than sole focus on psychotic symptoms’ (Gumley, Braehler, Laithwaite, MacBeth, & Gilbert, 2010). Given the evidence of the associations between FAR ability and psychosis and the impact that problems in this area can have on a person’s functioning, perhaps FAR, or more generally social cognition, should be assessed in all patients entering mental health services. Furthermore with the current evidence of the improvement in FAR ability after training, this may provide another way of aiding a person’s recovery from psychosis. In relation to recovery, it may also be useful to consider the impact of FAR difficulties within a therapeutic setting. If a person is unable to correctly interpret the emotion displayed by their therapist and misinterprets an empathetic stance as someone appearing disapproving or uninterested, this could impact their perceptions of feeling listened to and validated, which in turn may affect the therapeutic relationship.
References


http://schizophreniabulletin.oxfordjournals.org/content/early/2015/10/31/schbul.sbv150.abstract


Figure 1. Search process for non-affective psychosis vs. affective psychosis/bipolar disorder

Number of records identified through database searching: 4350

Number of records after duplicates removed: 2835

Number of records excluded on basis of title: 2715

Number of records screened (abstract/description): 120

Number of records excluded on abstract: 106

Number of full text reports screened for eligibility: 12

Number of full text reports excluded: 1 study used same sample as another

1 study did not have useable data for meta-analysis

Number of studies included in the meta-analysis: 10
Figure 2. Search process for active vs non-active psychosis

Number of records identified through database searching: 4350

Number of records after duplicates removed: 2835

Number of records screened (abstract/description): 120

Number of records excluded on basis of title: 2715

Number of full text reports screened for eligibility: 11

Number of records excluded on abstract: 109

Number of full text reports excluded:
- 1 FAR only measured at one time point
- 1 no useable data for meta-analysis

Number of studies included in the meta-analysis: 9
(6 repeated measures, 3 independent groups)
Figure 3. Search process for conversion to psychosis

Number of records identified through database searching: 4350

Number of records after duplicates removed: 2835

Number of records screened (abstract/description): 120

Number of records excluded on basis of title: 2715

Number of full text reports screened for eligibility: 5

Number of records excluded on abstract: 114

Number of full text reports excluded: 1 only measured psychosis-like symptoms

Number of studies included in the meta-analysis: 4
Figure 4. Search Process for FRT

Number of record identified through database searching: 3954

Number of records after duplicates removed: 2439

Number of records screened (abstract/description): 318

Number of records excluded on basis of title: 2121

Number of records excluded on abstract: 274

Number of full text reports screened for eligibility: 44

Number of full text reports excluded: 32
28 Less than 50% of intervention directly targeting FAR
3 Not specifically targeting FAR
1 not intervention trial

Number of studies included in the review: 13

Number of studies included in the meta-analysis: 8

Number of studies identified from reference section for inclusion: 1

Number of non-RCT trials: 5
<table>
<thead>
<tr>
<th>Trial</th>
<th>Psychosis Sample</th>
<th>Bipolar Sample</th>
<th>Control</th>
<th>FAR Measure</th>
<th>Symptom Measure</th>
<th>Location</th>
<th>Age (years)</th>
<th>Female n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thaler et al. (2013)</td>
<td>30 patients</td>
<td>48 patients</td>
<td>24 healthy controls</td>
<td>BLERT</td>
<td>SAPS, SANS, HAM-D</td>
<td>America</td>
<td>47.9 (10.8)P</td>
<td>34% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35.9 (13.2) BD</td>
<td>67% BD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36.1 (13.4) C</td>
<td>54% C</td>
</tr>
<tr>
<td>Lee et al. (2013)</td>
<td>38 patients</td>
<td>68 patients</td>
<td>36 healthy controls</td>
<td>Pictures of facial affect</td>
<td>BPRS, HAM-D, YMRS</td>
<td>Los Angeles</td>
<td>44.7 (9.9)P</td>
<td>33% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>43.9 (10.6) BP</td>
<td>75% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>41.4 (9.9) C</td>
<td>43% C</td>
</tr>
<tr>
<td>Addington et al (1998)</td>
<td>40 patients</td>
<td>40 patients</td>
<td>40 healthy controls</td>
<td>Pictures of Facial Affect</td>
<td>PANSS</td>
<td>Canada</td>
<td>32.6 (9.2)P</td>
<td>33% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38.5 (11.0) BP</td>
<td>75% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32.6 (11.3) C</td>
<td>43% C</td>
</tr>
<tr>
<td>Daros et al. (2014)</td>
<td>24 patients</td>
<td>16 patients</td>
<td>32 healthy controls</td>
<td>PEAT</td>
<td>PANSS, HAM-D, YMRS</td>
<td>Chicago</td>
<td>22.58 (5.69)P</td>
<td>21% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.63 (6.27) BP</td>
<td>44% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.78 (6.81) C</td>
<td>66% C</td>
</tr>
<tr>
<td>Yalcin-Siedentopf et al. (2014)</td>
<td>40 patients</td>
<td>57 patients</td>
<td>50 healthy controls</td>
<td>FEEL</td>
<td>PANSS, MADRS, YMRS</td>
<td>Innsbruck and Salzburg</td>
<td>40.3 (8.5)P</td>
<td>18% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>41.9 (11.7) BP</td>
<td>37% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38.8 (6.9) C</td>
<td>21% C</td>
</tr>
<tr>
<td>Edwards et al. (2001)</td>
<td>29 patients</td>
<td>23 patients</td>
<td>24 healthy controls</td>
<td>FACT</td>
<td>SAPS, SANS, HAM-D, BRMS</td>
<td>Melbourne</td>
<td>22.31 (4.12)P</td>
<td>24%P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22.43 (3.99) BP</td>
<td>39% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21.50 (4.09) C</td>
<td>37% C</td>
</tr>
<tr>
<td>Baez et al. (2013)</td>
<td>15 patients</td>
<td>15 patients</td>
<td>30 healthy controls</td>
<td>Emotional morphing</td>
<td>PANSS, MADRS, YMRS</td>
<td>Argentina</td>
<td>33.0 (7.9)P</td>
<td>20% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35.9 (11.8) BP</td>
<td>73% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34.3 (9.3) C</td>
<td>50% C</td>
</tr>
<tr>
<td>Goghari et al. (2013)</td>
<td>27 patients</td>
<td>16 patients</td>
<td>30 healthy controls</td>
<td>Pictures of facial affect</td>
<td>SAPS, SANS, BPRS</td>
<td>Minneapolis</td>
<td>38.9 (12.3)P</td>
<td>30% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>46.2 (11.3) BP</td>
<td>19% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48.8 (9.7) C</td>
<td>43% C</td>
</tr>
<tr>
<td>Vaskinn et al. (2007)</td>
<td>31 patients</td>
<td>21 patients</td>
<td>31 healthy controls</td>
<td>Pictures of facial affect</td>
<td>PANSS, YMRS</td>
<td>Norway</td>
<td>31.3 (9.5)P</td>
<td>35% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38.1 (9.3) BP</td>
<td>48% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30.7 (9.6) C</td>
<td>35% C</td>
</tr>
<tr>
<td>Rowland et al. (2013)</td>
<td>56 patients</td>
<td>33 patients</td>
<td>58 healthy controls</td>
<td>Pictures of facial affect</td>
<td>PANSS</td>
<td>Australia</td>
<td>44.57 (10.37)P</td>
<td>43% P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40.67 (11.27) BP</td>
<td>45% BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33.91 (12.24) C</td>
<td>50% C</td>
</tr>
</tbody>
</table>
### Table 2. Study characteristics of active vs non-active psychosis

Repeated measures studies

<table>
<thead>
<tr>
<th>Trial</th>
<th>Sample</th>
<th>Control</th>
<th>FAR Measure</th>
<th>Symptom Measure</th>
<th>Location</th>
<th>Age of onset mean (s.d)</th>
<th>Age (years)mean (s.d)</th>
<th>Female n (%)</th>
<th>Follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lewis &amp; Garver (1995)</td>
<td>18 patients</td>
<td>10 matched controls</td>
<td>Pictures of Facial Affect</td>
<td>BPRS</td>
<td>Dallas</td>
<td>22.5 (3.2)</td>
<td>38.9 (6.5)P 38.5 (7.2)C</td>
<td>2 weeks</td>
<td></td>
</tr>
<tr>
<td>Maat et al. (2015)</td>
<td>521 patients</td>
<td>312 healthy controls</td>
<td>Degraded Facial Affect Recognition</td>
<td>PANSS</td>
<td>Netherlands</td>
<td>22.57 (6.8)</td>
<td>27.34(7.33)P 30.12(10.73)C</td>
<td>23%P 50%C</td>
<td>3 years</td>
</tr>
<tr>
<td>Addington et al. (2006)</td>
<td>50 patients</td>
<td>55 matched controls</td>
<td>FEIT</td>
<td>PANSS</td>
<td>Canada</td>
<td>25.1 (8.01)P 21.7(6.05)C</td>
<td>30.12(10.73)C</td>
<td>40%P 40%C</td>
<td>1 year</td>
</tr>
<tr>
<td>Addington et al. (1998)</td>
<td>40 patients</td>
<td>40 healthy controls</td>
<td>Pictures of Facial Affect</td>
<td>PANSS</td>
<td>Canada</td>
<td>22.5 (7.2)</td>
<td>32.6 (9.2)P 32.6(11.3)C</td>
<td>33%P 43%C</td>
<td>3 months</td>
</tr>
<tr>
<td>Wolwer et al. (1996)</td>
<td>32 patients</td>
<td>21 healthy controls</td>
<td>Pictures of Facial Affect</td>
<td>BPRS</td>
<td>Dusseldorf</td>
<td>31.7(10.6)P 34.2(10)C</td>
<td>30.12(10.73)C</td>
<td>31%P 29%C</td>
<td>4 weeks</td>
</tr>
<tr>
<td>Daros et al. (2014)</td>
<td>24 patients</td>
<td>32 healthy controls</td>
<td>PEAT</td>
<td>PANSS</td>
<td>Chicago</td>
<td>22.58(5.69)P 25.78(6.81)C</td>
<td>21%P 66%C</td>
<td>7 weeks</td>
<td></td>
</tr>
</tbody>
</table>

Independent groups

<table>
<thead>
<tr>
<th>Trial</th>
<th>Sample</th>
<th>Control</th>
<th>FAR Measure</th>
<th>Symptom Measure</th>
<th>Location</th>
<th>Age of onset mean (s.d)</th>
<th>Age (years)mean (s.d)</th>
<th>Female n (%)</th>
<th>Follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behere et al (2001)</td>
<td>63 patients</td>
<td>45 matched controls</td>
<td>TRENDS</td>
<td>SAPS</td>
<td>Bangalore</td>
<td>28.3 (7.5)P 27.8 (7.3)FRS 28.7 (7.8) noFRS</td>
<td>29.8(8.1)P 29.0(8.9)C 30.5(8.7)FRS 29.3(7.7)noFRS</td>
<td>54%P 40%C 62%FRS 49%noFRS</td>
<td></td>
</tr>
<tr>
<td>Penn et al. (2000)</td>
<td>74 patients</td>
<td>40 healthy controls</td>
<td>FEIT</td>
<td>BPRS</td>
<td>Louisiana</td>
<td>16.9(9.3)EC 13.4(8.5)AC</td>
<td>40.15(8.09)EC 36.4(10.02)AC 32.33(10.79)C</td>
<td>44%EC 14%AC 53%C</td>
<td></td>
</tr>
<tr>
<td>Pan et al. (2009)</td>
<td>73 patients</td>
<td>40 healthy controls</td>
<td>DAVANA-2</td>
<td>PANSS</td>
<td>Taiwan</td>
<td>24.5(8.31)A 21.97(6.3)S</td>
<td>30.6(7.43)A 32.76(8.14)S 33.75(10.64)C</td>
<td>50%A 45.5%S 62.5%C</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Study characteristics of conversion to psychosis

<table>
<thead>
<tr>
<th>Trial</th>
<th>Sample</th>
<th>Conversion to Psychosis</th>
<th>Control</th>
<th>FAR Measure</th>
<th>Symptom Measure</th>
<th>Location</th>
<th>Age (years)mean (s.d)</th>
<th>Female n (%)</th>
<th>Follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piskulic et al. (2016)</td>
<td>675 CHR</td>
<td>86 patients</td>
<td>264 healthy controls</td>
<td>ER40</td>
<td>SIPS</td>
<td>North America</td>
<td>18.5(4.23)CHR 19.73(4.67)C</td>
<td>42%CHR</td>
<td>50% C</td>
</tr>
<tr>
<td>Corcoran et al (2015)</td>
<td>49 CHR</td>
<td>7 patients</td>
<td>31 healthy controls</td>
<td>ER40</td>
<td>EMODIFF</td>
<td>New York</td>
<td>20.7(3.5)CHR- 20.0(5.2)CHR+ 21.4(3.1)C</td>
<td>24%CHR-</td>
<td>43%CHR+</td>
</tr>
<tr>
<td>Addington et al. (2012)</td>
<td>172 CHR</td>
<td>29 patients</td>
<td>100 help seeking controls</td>
<td>FEIT</td>
<td>SIPS</td>
<td>Toronto, North Carolina, Connecticut</td>
<td>19.4(3.9)CHR 19.4(3.9)C</td>
<td>43%CHR</td>
<td>44% C</td>
</tr>
<tr>
<td>Allott et al (2014)</td>
<td>37 CHR</td>
<td>11 patients</td>
<td>No control</td>
<td>PFA</td>
<td>PANSS</td>
<td>Vienna</td>
<td>16.2 (1.7)</td>
<td>67.6%</td>
<td>1 year</td>
</tr>
</tbody>
</table>
Table 4. Study Characteristics

<table>
<thead>
<tr>
<th>Trial</th>
<th>Treatments</th>
<th>Number randomized</th>
<th>Maximum duration of treatment in sessions</th>
<th>FAR measures</th>
<th>Additional measures</th>
<th>Inpatients</th>
<th>Age (years)mean (s.d)</th>
<th>Female n (%)</th>
<th>Follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penn, D. L., Combs, D. (2000)</td>
<td>Monetary Reinforcement and Facial Feedback;</td>
<td>9</td>
<td>1</td>
<td>Face Emotion Identification Test (FEIT); Facial Emotional Discrimination Task (FEDT)</td>
<td></td>
<td>100%</td>
<td>38.3 (6.04)</td>
<td>44%</td>
<td>1 week</td>
</tr>
<tr>
<td></td>
<td>Monetary Feedback only</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40.42 (6.08)</td>
<td>42%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9 Facial Feedback only</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>39.1 (8.3)</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>repeated exposure/active control</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>41.5 (12)</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Russell T.A., Green M.J., Simpson I., Coltheart M. (2008)</td>
<td>Micro-expression training tool (METT)</td>
<td>26</td>
<td>1</td>
<td>EMT (emotion matching task) pre and post</td>
<td></td>
<td>0%</td>
<td>40 (10)</td>
<td>35%</td>
<td>1 week</td>
</tr>
<tr>
<td></td>
<td>repeated exposure/active control</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44 (9)</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>Study (Year)</td>
<td>Intervention/Method</td>
<td>Sample Size</td>
<td>Duration (weeks)</td>
<td>PANSS Measure</td>
<td>Social and Occupational Functioning Assessment (SOFAS)</td>
<td>Completion Rate</td>
<td>Improvement</td>
<td>Adverse Events</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>---------------</td>
<td>--------------------------------------------------------</td>
<td>----------------</td>
<td>------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Wolwer W., Frommann N. (2011)</td>
<td>Training of Affect Recognition (TAR)</td>
<td>20</td>
<td>12</td>
<td>Pictures of Facial Affect</td>
<td>Social and Occupational Functioning Assessment (SOFAS); PANSS pre and post</td>
<td>100%</td>
<td>36.7 (13.1)</td>
<td>32%</td>
<td>none</td>
</tr>
<tr>
<td>Combs D.R., Tosheva A., Penn D.L., Basso M.R., Wanner J.L., Laib K. (2008)</td>
<td>Attentional Shaping</td>
<td>20</td>
<td>1</td>
<td>Face Emotion Identification Test (FEIT); Bell-Lysaker Emotion Recognition Test (BLERT)</td>
<td>100%</td>
<td>38.7 (13.7)</td>
<td>35%</td>
<td>1 week</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>n</td>
<td>Duration</td>
<td>PANSS Pre</td>
<td>PANSS Post</td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------------------------</td>
<td>---</td>
<td>-------------</td>
<td>-----------</td>
<td>------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habel, Ute, Koch, Kathrin, Kellermann, Thilo, Reske, Martina, Frommann, Nicole, Wolwer, Wolfgang, Zilles, Karl, Shah, N. Jon, Schneider, Frank (2010)</td>
<td>Monetary Reinforcement repeated exposure/active control</td>
<td>20</td>
<td></td>
<td>31.4 (7.8)</td>
<td>0%</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luckhaus, Christian, Frommann, Nicole, Stroth, Sanna, Brinkmeyer, Jurgen, Wolwer, Wolfgang (2013)</td>
<td>Training of Affect Recognition (TAR) TAU</td>
<td>10</td>
<td>2 months</td>
<td>35.3 (8.2)</td>
<td>0%</td>
<td>2 months</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Consistency of FAR and Psychosis- Heterogeneity of included studies

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study</th>
<th>Chi-square (df) , p</th>
<th>I2</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAR in psychosis vs FAR in healthy</td>
<td>Kohler (2009)</td>
<td>Q(76)=295.7, p&lt;0.001</td>
<td>74%</td>
</tr>
<tr>
<td>FAR correlation with positive symptoms</td>
<td>Ventura (2013)</td>
<td>Q(22)=34.23, p=0.05</td>
<td>36%</td>
</tr>
<tr>
<td>FAR correlation with negative symptoms</td>
<td>Ventura (2013)</td>
<td>Q(59)=132.35, p=0.00</td>
<td>55%</td>
</tr>
<tr>
<td>FAR correlation with reality distortion</td>
<td>Ventura (2013)</td>
<td>Q(18)=45.70, p=0.00</td>
<td>60%</td>
</tr>
<tr>
<td>FAR correlation with disorganisation</td>
<td>Ventura (2013)</td>
<td>Q(22)=39.21, p=0.01</td>
<td>44%</td>
</tr>
<tr>
<td>FAR in non-affective vs affective psychosis</td>
<td>Current</td>
<td>Q= 491.08, p= 0.00</td>
<td>98%</td>
</tr>
<tr>
<td>FAR in active vs non-active psychosis</td>
<td>Current</td>
<td>Q=2.56, p= 0.96</td>
<td>0%</td>
</tr>
<tr>
<td>FAR in CHR conversion to psychosis</td>
<td>Current</td>
<td>Q=1.93, p= 0.59</td>
<td>0%</td>
</tr>
<tr>
<td>FAR in CHR vs control</td>
<td>Current</td>
<td>Q=1.61, p=0.45</td>
<td>0%</td>
</tr>
<tr>
<td>FRT effect post intervention</td>
<td>Current</td>
<td>Q=20.15, p=0.01</td>
<td>65%</td>
</tr>
<tr>
<td>FRT effect 1 week follow up</td>
<td>Current</td>
<td>Q=7.21, p=0.03</td>
<td>72%</td>
</tr>
<tr>
<td>FRT effect vs active control</td>
<td>Current</td>
<td>Q=16.57, p=0.00</td>
<td>76%</td>
</tr>
<tr>
<td>FRT effect on negative symptoms</td>
<td>Current</td>
<td>Q=1.27, p=0.74</td>
<td>0%</td>
</tr>
<tr>
<td>FRT effect on positive symptoms</td>
<td>Current</td>
<td>Q=0.55, p=0.76</td>
<td>0%</td>
</tr>
<tr>
<td>FRT effect on general psychopathology</td>
<td>Current</td>
<td>Q=4.55, p=0.10</td>
<td>56%</td>
</tr>
</tbody>
</table>

Figure 5. Specificity of FAR and Psychosis- Comparison FAR ability in non-affective psychosis and affective psychosis/bipolar disorder
Figure 6(a). Temporality of FAR and Psychosis- Comparison of CHR with control subjects

ES (95% CI) % Weight
-0.26 (-0.41, -0.12) 3.59
-0.35 (-0.49, -0.21) 96.40
-0.23 (-3.14, 2.34) 0.01
-0.35 (-0.53, -0.17) 100.00

Figure 6(b). Temporality of FAR and Psychosis- Comparison of FAR ability in CHR who convert to psychosis with those who do not convert

ES (95% CI) % Weight
-0.04 (-1.19, -0.08) 87.16
-0.32 (-1.71, 0.07) 9.33
0.01 (0.71, 0.71) 2.99
-0.73 (-4.35, 2.76) 0.12
-0.07 (-0.19, 0.06) 100.00

Figure 7. Temporality of FAR and Psychosis- Comparison of FAR ability in active and non-active psychosis

ES (95% CI) % Weight
-0.06 (-0.57, 0.46) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.28, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.22 (0.32, 0.08) 0.01
0.52 (0.79, 1.16) 3.86
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.28, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
0.06 (-0.57, 0.45) 15.29
0.14 (0.59, 0.87) 7.48
0.23 (1.17, 1.39) 4.59
0.02 (0.32, 0.03) 63.96
0.52 (0.69, 1.93) 1.98
0.27 (1.17, 3.86) 3.86
-0.21 (-1.74, 1.34) 0.01
0.01 (-0.79, 1.11) 1.33
0.14 (-0.74, 1.27) 1.33
0.23 (1.07, 1.33) 4.03
Figure 8(a). Experiment Criterion- The effect of facial affect recognition training on recognition of facial affect at post-intervention

Figure 8(b). Experiment Criterion- The effect of facial affect recognition training on recognition of facial affect at 1-week follow up

Figure 8(c). Experiment Criterion- Facial affect recognition training versus active control groups at post-intervention.
Figure 9(a). Experiment Criterion- The effect of facial affect recognition training on negative symptoms

Figure 9(b). Experiment Criterion- The effect of facial affect recognition training on positive symptoms

Figure 9(c). Experiment Criterion- The effect of facial affect recognition training on general psychopathology
### Table 6. Summary of Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of included studies</th>
<th>No. of included studies</th>
<th>Psychosis, n</th>
<th>Bipolar, n</th>
<th>Hedges g (95% CI)</th>
<th>Heterogeneity</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAR ability in non-affective psychosis vs affective psychosis</td>
<td>10</td>
<td></td>
<td>330</td>
<td>337</td>
<td>-0.62 (-1.21, -0.03)</td>
<td>P= 98%</td>
<td>Very low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of included studies</th>
<th>No. of included studies</th>
<th>CHR, n</th>
<th>Healthy control, n</th>
<th>Cohens d (95% CI)</th>
<th>Heterogeneity</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAR ability in CHR vs healthy control</td>
<td>3</td>
<td></td>
<td>896</td>
<td>395</td>
<td>-0.35 (-0.38, -0.32)</td>
<td>P= 0%</td>
<td>Low</td>
</tr>
<tr>
<td>FAR ability in converters to psychosis vs non-converters</td>
<td>4</td>
<td></td>
<td>933 (133 converters)</td>
<td>395</td>
<td>-0.07 (-0.19, 0.06)</td>
<td>P= 0%</td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of included studies</th>
<th>No. of included studies</th>
<th>Active psychosis, n</th>
<th>Non-active psychosis, n</th>
<th>Hedges g (95% CI)</th>
<th>Heterogeneity</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAR ability in active vs non-active psychosis</td>
<td>9</td>
<td></td>
<td>386</td>
<td>394</td>
<td>0.02 (-0.18, 0.22)</td>
<td>P= 0%</td>
<td>Very low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>No. of treatment sessions</th>
<th>No. of included studies</th>
<th>Intervention, n</th>
<th>Control, n</th>
<th>Hedges g (95% CI)</th>
<th>Heterogeneity</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAR improvement after intervention</td>
<td>1-12</td>
<td>8</td>
<td>152</td>
<td>148</td>
<td>1.35 (0.89, 1.80)</td>
<td>P= 65%</td>
<td>Low</td>
</tr>
<tr>
<td>FAR improvement at follow up</td>
<td>1</td>
<td>3</td>
<td>64</td>
<td>44</td>
<td>1.46 (0.61, 2.32)</td>
<td>P= 72%</td>
<td>Very low</td>
</tr>
<tr>
<td>FAR improvement vs active control group</td>
<td>1-12</td>
<td>5</td>
<td>112</td>
<td>86</td>
<td>1.60 (0.91, 2.28)</td>
<td>P= 76%</td>
<td>Very low</td>
</tr>
<tr>
<td>Improvement in negative symptoms</td>
<td>12</td>
<td>4</td>
<td>78</td>
<td>95</td>
<td>-0.11 (-0.41, 0.20)</td>
<td>P= 0%</td>
<td>Very low</td>
</tr>
<tr>
<td>Improvement in positive symptoms</td>
<td>12</td>
<td>3</td>
<td>58</td>
<td>77</td>
<td>0.10 (-0.25, 0.45)</td>
<td>P= 0%</td>
<td>Very low</td>
</tr>
</tbody>
</table>
Improvement in general psychopathology  | 12 | 3 | 58 | 77 | 0.12 (-0.44, 0.69) | P=56% | Very low

**Table 7(a). Quality Analysis of Non-RCT studies- Non-affective psychosis vs affective psychosis**

<table>
<thead>
<tr>
<th>Study</th>
<th>Unbiased selection of cohort</th>
<th>selection minimises baseline differences in prognostic factors</th>
<th>Sample size justification report</th>
<th>Sufficient power</th>
<th>adequate description of the cohort</th>
<th>Validated method for measuring symptoms</th>
<th>Validated method for measuring facial affect recognition</th>
<th>Outcome assessment blind to exposure</th>
<th>Analytic methods appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thaler (2013)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Lee (2013)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Addington (1998)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Daros (2014)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
</tr>
<tr>
<td>Yalcin-S (2014)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Edwards (2001)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Baez (2013)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Goghari (2013)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Vaskinn (2007)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Rowland (2013)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
</tbody>
</table>
Table 7(b) Quality Analysis of Non-RCT studies- Active vs Non-active Psychosis

<table>
<thead>
<tr>
<th>Study</th>
<th>Unbiased selection of cohort</th>
<th>selection minimises baseline differences in prognostic factors</th>
<th>Sample size justification report</th>
<th>Sufficient power</th>
<th>adequate description of the cohort</th>
<th>Validated method for measuring symptoms</th>
<th>Validated method for measuring facial affect recognition</th>
<th>Outcome assessment blind to exposure</th>
<th>Analytic methods appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lewis (1995)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Maat (2015)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Addington (2006)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>Partially</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
</tr>
<tr>
<td>Addington (1998)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Wowler (1996)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Daros (2014)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
</tr>
<tr>
<td>Behere (2011)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
</tr>
<tr>
<td>Pan (2009)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Penn (2000)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
</tbody>
</table>

Table 7(c) Quality Analysis of Non-RCT studies- conversion to psychosis

<table>
<thead>
<tr>
<th>Study</th>
<th>Unbiased selection of cohort</th>
<th>selection minimises baseline differences in prognostic factors</th>
<th>Sample size justification report</th>
<th>Sufficient power</th>
<th>adequate description of the cohort</th>
<th>Validated method for measuring symptoms</th>
<th>Validated method for measuring facial affect recognition</th>
<th>Outcome assessment blind to exposure</th>
<th>Analytic methods appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piskulic (2016)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>Partially</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Yes</td>
</tr>
<tr>
<td>Addington (2012)</td>
<td>Partially</td>
<td>Yes</td>
<td>No</td>
<td>Partially</td>
<td>Yes</td>
<td>Partially</td>
<td>Yes</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Allott (2014)</td>
<td>Partially</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
<tr>
<td>Corcoran (2015)</td>
<td>Partially</td>
<td>Partially</td>
<td>No</td>
<td>No</td>
<td>Partially</td>
<td>Partially</td>
<td>Partially</td>
<td>Can’t tell</td>
<td>Partially</td>
</tr>
</tbody>
</table>
### Table 7(d) Risk of Bias Table for FRT Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data addressed (attrition bias)</th>
<th>Incomplete outcome data addressed (attrition bias) (Follow-up)</th>
<th>Selective reporting (reporting bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combs (2008)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Habel (2010)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>N/A</td>
<td>Low</td>
</tr>
<tr>
<td>Luckhaus (2013)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Penn (2000)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Sachs (2012)</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>N/A</td>
<td>Low</td>
</tr>
<tr>
<td>Russell (2008)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Wolwer (2005)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>N/A</td>
<td>Low</td>
</tr>
<tr>
<td>Wolwer (2011)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

### Table 8. GRADE Assessment of Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Quality</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication Bias</th>
<th>Other factors</th>
<th>Overall</th>
<th>Included Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAR ability and psychosis</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Low</td>
<td>See Kohler et al. (2009)</td>
</tr>
<tr>
<td>FAR ability and positive symptoms</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>Low</td>
<td>See Ventura et al. (2013)</td>
</tr>
<tr>
<td>FAR ability and negative symptoms</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>Very low</td>
<td>See Ventura et al. (2013)</td>
</tr>
<tr>
<td>FAR ability and reality distortion</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>Very low</td>
<td>See Ventura et al. (2013)</td>
</tr>
<tr>
<td>FAR ability and conceptual disorganisation</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>Very low</td>
<td>See Ventura et al. (2013)</td>
</tr>
<tr>
<td>FAR ability in CHR vs healthy control</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>Low</td>
<td>Piskulic (2016), Corcoran (2015), Addington (2012), Allott</td>
</tr>
<tr>
<td>Outcome Description</td>
<td>Rating</td>
<td>Count (Minimum)</td>
<td>Count (Maximum)</td>
<td>Rating</td>
<td>References</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>--------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAR ability in those who convert to psychosis</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAR ability in active vs non-active psychosis</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAR improvement after intervention</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAR improvement at follow up</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAR improvement vs active control group</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement in negative symptoms</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement in positive symptoms</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement in general psychopathology</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>Very low</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For assessment of outcome quality, a downgrade of 1 point was made if >50% of studies contributing to that outcome had at least one ‘high risk’ rating according to the Cochrane Risk of Bias assessment or at least one ‘no’ on the AHRQ and 2 points if >50% of studies has at least two ratings of ‘high risk’ or ‘no’. For inconsistency, a study was downgraded by 1 point if the I² statistic was >40% in the context of an unclear direction of effect or >75% in the context of a clear direction of effect. If the I² statistic was >75% in the context of no clear direction of effect, a downgrade of 2 points was made. For imprecision, an outcome was downgraded if “a recommendation or clinical course of action would differ if the upper versus the lower boundary of the CI represented the truth” (Guyatt et al., 2011). A downgrade was made for publication bias for outcomes with less than 10 studies as it therefore cannot be assessed (Ioannidis & Trikalinos, 2007). An outcome was upgraded by 1 point if a very large effect size was found (Higgins et al., 2011).
Appendix 1. AMSTAR Quality Assessment

AMSTAR – a measurement tool to assess the methodological quality of systematic reviews.

1. Was an ‘a priori’ design provided?
   The research question and inclusion criteria should be established before the conduct of the review.
   Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a “yes.”
   ✔ Yes
   □ No
   □ Can’t answer
   □ Not applicable

2. Was there duplicate study selection and data extraction?
   There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.
   Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other’s work.
   □ Yes
   ✔ No
   □ Can’t answer
   □ Not applicable

3. Was a comprehensive literature search performed?
   At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.
   Note: If at least 2 sources + one supplementary strategy used, select “yes” (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).
   ✔ Yes
   □ No
   □ Can’t answer
   □ Not applicable

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
   The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.
   Note: If review indicates that there was a search for “grey literature” or “unpublished literature,” indicate “yes.” SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.
   ✔ Yes
   □ No
   □ Can’t answer
   □ Not applicable

5. Was a list of studies (included and excluded) provided?
   A list of included and excluded studies should be provided.
   Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select “no.”
6. Were the characteristics of the included studies provided?
In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.
Note: Acceptable if not in table format as long as they are described as above.

7. Was the scientific quality of the included studies assessed and documented?
'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.
Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study (“low” or “high” is fine, as long as it is clear which studies scored “low” and which scored “high”; a summary score/range for all studies is not acceptable).

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.
Note: Might say something such as “the results should be interpreted with caution due to poor quality of included studies.” Cannot score “yes” for this question if scored “no” for question 7.

9. Were the methods used to combine the findings of studies appropriate?
For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I²). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).
Note: Indicate “yes” if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.

10. Was the likelihood of publication bias assessed?
An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).

Note: If no test values or funnel plot included, score “no”. Score “yes” if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

- Yes
- No
- Can't answer
- Not applicable

11. Was the conflict of interest included?
Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

Note: To get a “yes,” must indicate source of funding or support for the systematic review AND for each of the included studies.

- Yes
- No
- Can't answer
- Not applicable

Additional notes (in italics) made by Michelle Weir, Julia Worswick, and Carolyn Wayne based on conversations with Bev Shea and/or Jeremy Grimshaw in June and October 2008 and July and September 2010.
Appendix 2. Quality Assessment Tool.

Quality assessment of observational studies


General instructions: Grade each criterion as “Yes,” “No,” “Partially,” or “Can’t tell.” Factors to consider when making an assessment are listed under each criterion.

1. Unbiased selection of the cohort?

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The participants in the study are likely to be representative of the target population. The recruitment strategy is clearly described and less likely to introduce bias.</td>
</tr>
<tr>
<td>No</td>
<td>The sample is not likely to be representative of the target population. The recruitment strategy is not described and/or is likely to introduce bias.</td>
</tr>
<tr>
<td>Partially</td>
<td>The participants are less likely to be representative of the target population. The recruitment strategy is somewhat likely to introduce bias.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

2. Selection minimizes baseline differences in prognostic factors

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The selection of a comparison group was appropriate and the group are unlikely to differ on factors related to the outcome (besides antisocial factors). Or the authors indicated that 80-100% of confounders (age, sex, education, IQ, ethnicity) were controlled for in the design (matching) or in the analysis (propensity scores).</td>
</tr>
<tr>
<td>No</td>
<td>There were clear differences in confounding variables between groups of which &lt;60% were controlled for in the design or analysis.</td>
</tr>
<tr>
<td>Partially</td>
<td>The group differed on confounding variables and/or some (60-79%) of which were controlled for in the design or analysis.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>
3. Sample size justification reported

<table>
<thead>
<tr>
<th>Yes</th>
<th>Power calculation and effect size estimation was clearly reported.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No evidence or justification of sample size.</td>
</tr>
<tr>
<td>Partially</td>
<td>Limited evidence or justification of sample size.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

4. Sufficient power

G*Power 3.1.6 (Faul, Erdfelder, Lang & Bucher, 2007) was used to calculate sample sizes required for sufficient power. For correlational analyses it is necessary to recruit 21 participants to detect a large effect size \((r=0.5)\), 62 to detect a moderate effect size \((r=0.3)\) and 614 participants to detect a small effect size \((r=0.1)\) with the statistical power of 0.8 at an alpha level of 0.05. For differences between groups it is necessary to recruit 21 to each group to detect a large effect size \((d=0.8)\), 51 to detect a moderate effect size \((d=0.5)\) and 310 in each group to detect a small effect size \((d=0.2)\) with the statistical power of 0.8 at an alpha level of 0.05.

<table>
<thead>
<tr>
<th>Yes</th>
<th>The study has a sample size large enough to detect small to moderate group differences ((d=0.2-0.5)) or correlations ((r=0.1-0.3)) with the statistical power of 0.8 at an alpha level of 0.05.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The study has a sample size large enough to detect large to very large differences or correlations with the statistical power of 0.8 at an alpha level of 0.05.</td>
</tr>
<tr>
<td>Partially</td>
<td>The study has a sample size large enough to detect moderate to large group differences ((d=0.5-0.8)) or correlations ((r=0.3-0.5)) with the statistical power of 0.8 at an alpha level of 0.05.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

5. Adequate description of the cohort?

<table>
<thead>
<tr>
<th>Yes</th>
<th>The cohort is clearly ((&gt;4)) specified and defined in terms of baseline demographics (age, gender, ethnicity, setting, IQ)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>The sample is poorly described in terms of key baseline demographics (&lt;2).</td>
</tr>
<tr>
<td>Partially</td>
<td>The cohort is less well (&lt;3) specified and defined in terms of baseline demographics (age, gender, ethnicity, setting, IQ/educational achievement)?</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>
6. Validated method for measuring facial affect recognition deficits?

<table>
<thead>
<tr>
<th>Response</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The psychometric properties of the outcome measure are clearly reported and are valid and reliable in the study population.</td>
</tr>
<tr>
<td>No</td>
<td>The outcome measure has not been described in any detail and/or has not undergone psychometric evaluation.</td>
</tr>
<tr>
<td>Partially</td>
<td>The outcome measure is described less clearly and psychometric properties have not been described and/or the measure has not been validated in this population.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

7. Outcome assessment blind to exposure?

<table>
<thead>
<tr>
<th>Response</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The study investigators who assessed outcomes were blind to the aggressive history status of the participants. Participants were blind to the research question.</td>
</tr>
<tr>
<td>No</td>
<td>The study investigators who assessed outcomes were not blind to the aggressive history status of the participants. The participants were not blind to the research question.</td>
</tr>
<tr>
<td>Partially</td>
<td>Either the study investigators who assessed outcomes were blind to the aggressive history status of the participants or the participants were blind to the research question.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

8. Analytic methods appropriate?\(^5\)

<table>
<thead>
<tr>
<th>Response</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The method of statistical analysis was appropriate to the research question being asked. Confidence intervals, p-values and effect sizes are reported.</td>
</tr>
<tr>
<td>No</td>
<td>The method of analysis is not appropriate to the research question and does not provide meaningful results. The study investigators who assessed outcomes were not blind to the antisocial status of the participants. The participants were not blind to the research question.</td>
</tr>
<tr>
<td>Partially</td>
<td>The analysis is appropriate however the findings are not reported in sufficient detail.</td>
</tr>
<tr>
<td>Can’t tell</td>
<td>The study has not reported this information or it is not applicable in this case.</td>
</tr>
</tbody>
</table>

\(^5\) As recommended in Discovering Statistics Using IBM SPSS Statistics(Field, 2013).
Research Portfolio References


Pan, Y.-J., Chen, S.-H., Chen, W. J., & Liu, S.-K. (2009). Affect recognition as an independent social function determinant in schizophrenia. *Comprehensive psychiatry, 50*(5), 443-452. @ 0010-0440X.


Research Portfolio Appendix 1. Instructions for Contributors

Psychological Review

Prior to submission, please carefully read and follow the submission guidelines detailed below. Manuscripts that do not conform to the submission guidelines may be returned without review.

Submission

Submit manuscripts electronically (.doc or .pdf format) to:

Keith J. Holyoak, Editor
Department of Psychology
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90095-1563

General correspondence may be directed to the Editor's Office.

Do not submit manuscripts to the Editor's email address.

All submissions should be clear and readable. An unusual typeface is acceptable only if it is clear and legible.

In addition to addresses and phone numbers, please supply electronic mail addresses and fax numbers, if available, for potential use by the editorial office and later by the production office.

Psychological Review® is now using a software system to screen submitted content for similarity with other published content. The system compares the initial version of each submitted manuscript against a database of 40+ million scholarly documents, as well as content appearing on the open web. This allows APA to check submissions for potential overlap with material previously published in scholarly journals (e.g., lifted or republished material).

Masked Review Policy

Open (i.e., unmasked) review is the default for this journal, though masked review is an option. If you choose masked review, include authors' names and affiliations only in the cover letter for the manuscript.

Authors who choose masked review should make every effort to see that the manuscript itself contains no clues to their identities.

Length

There is no upper bound on the length of Psychological Review articles. However, authors who submit papers with texts longer than 25,000 words will be asked to justify the need for their length.

Submissions must be under 5 MB in total size.

Manuscript Preparation
Prepare manuscripts according to the Publication Manual of the American Psychological Association (6th edition). Manuscripts may be copyedited for bias-free language (see Chapter 3 of the Publication Manual).

Review APA's Checklist for Manuscript Submission before submitting your article.

Double-space all copy. Other formatting instructions, as well as instructions on preparing tables, figures, references, metrics, and abstracts, appear in the Manual. Additional guidance on APA Style is available on the APA Style website.

Below are additional instructions regarding the preparation of display equations, computer code, and tables.

Display Equations

We strongly encourage you to use MathType (third-party software) or Equation Editor 3.0 (built into pre-2007 versions of Word) to construct your equations, rather than the equation support that is built into Word 2007 and Word 2010. Equations composed with the built-in Word 2007/Word 2010 equation support are converted to low-resolution graphics when they enter the production process and must be rekeyed by the typesetter, which may introduce errors.

To construct your equations with MathType or Equation Editor 3.0:

Go to the Text section of the Insert tab and select Object.

Select MathType or Equation Editor 3.0 in the drop-down menu.

If you have an equation that has already been produced using Microsoft Word 2007 or 2010 and you have access to the full version of MathType 6.5 or later, you can convert this equation to MathType by clicking on MathType Insert Equation. Copy the equation from Microsoft Word and paste it into the MathType box. Verify that your equation is correct, click File, and then click Update. Your equation has now been inserted into your Word file as a MathType Equation.

Use Equation Editor 3.0 or MathType only for equations or for formulas that cannot be produced as Word text using the Times or Symbol font.

Computer Code

Authors of accepted articles who report new computer simulations of models, or new data-analysis software, are required to provide the code as online supplemental material ("additional content") at the time of final manuscript submission. It is important to include adequate documentation so that the code can be downloaded and used by other researchers.

Because altering computer code in any way (e.g., indents, line spacing, line breaks, page breaks) during the typesetting process could alter its meaning, we treat computer code differently from the rest of your article in our production process. To that end, we request separate files for computer code.

In Online Supplemental Material

If you are interested in including supplemental material with your article, please state your intentions in the cover letter that accompanies your submission to the Editor. If you wish the material to be made available to reviewers, then include it with your initial submission.
For more information, visit Supplementing Your Article With Online Material.

In the Text of the Article
If you would like to include code in the text of your published manuscript, please submit a separate file with your code exactly as you want it to appear, using Courier New font with a type size of 8 points. We will make an image of each segment of code in your article that exceeds 40 characters in length. (Shorter snippets of code that appear in text will be typeset in Courier New and run in with the rest of the text.) If an appendix contains a mix of code and explanatory text, please submit a file that contains the entire appendix, with the code keyed in 8-point Courier New.

Tables
Use Word's Insert Table function when you create tables. Using spaces or tabs in your table will create problems when the table is typeset and may result in errors.

Submitting Supplemental Materials
APA can place supplemental materials online, available via the published article in the PsycARTICLES® database. Please see Supplementing Your Article With Online Material for more details.

Abstract and Keywords
All manuscripts must include an abstract containing a maximum of 250 words typed on a separate page. After the abstract, please supply up to five keywords or brief phrases.

References
List references in alphabetical order. Each listed reference should be cited in text, and each text citation should be listed in the References section.

Examples of basic reference formats:

Journal Article:

Authored Book:

Chapter in an Edited Book:

Figures
Graphics files are welcome if supplied as Tiff or EPS files. Multipanel figures (i.e., figures with parts labeled a, b, c, d, etc.) should be assembled into one file.

The minimum line weight for line art is 0.5 point for optimal printing.
For more information about acceptable resolutions, fonts, sizing, and other figure issues, please see the general guidelines.

When possible, please place symbol legends below the figure instead of to the side.

APA offers authors the option to publish their figures online in color without the costs associated with print publication of color figures.

The same caption will appear on both the online (color) and print (black and white) versions. To ensure that the figure can be understood in both formats, authors should add alternative wording (e.g., "the red (dark gray) bars represent") as needed.

For authors who prefer their figures to be published in color both in print and online, original color figures can be printed in color at the editor's and publisher's discretion provided the author agrees to pay:

$900 for one figure

An additional $600 for the second figure

An additional $450 for each subsequent figure

Permissions

Authors of accepted papers must obtain and provide to the editor on final acceptance all necessary permissions to reproduce in print and electronic form any copyrighted work, including test materials (or portions thereof), photographs, and other graphic images (including those used as stimuli in experiments).

On advice of counsel, APA may decline to publish any image whose copyright status is unknown.

Download Permissions Alert Form (PDF, 13KB)

Publication Policies

APA policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications.

See also APA Journals® Internet Posting Guidelines.

APA requires authors to reveal any possible conflict of interest in the conduct and reporting of research (e.g., financial interests in a test or procedure, funding by pharmaceutical companies for drug research).

Download Disclosure of Interests Form (PDF, 38KB)

In light of changing patterns of scientific knowledge dissemination, APA requires authors to provide information on prior dissemination of the data and narrative interpretations of the data/research appearing in the manuscript (e.g., if some or all were presented at a conference or meeting, posted on a listserv, shared on a website, including academic social networks like ResearchGate, etc.). This information (2–4 sentences) must be provided as part of the Author Note.

Authors of accepted manuscripts are required to transfer the copyright to APA.
Ethical Principles

It is a violation of APA Ethical Principles to publish "as original data, data that have been previously published" (Standard 8.13).

In addition, APA Ethical Principles specify that "after research results are published, psychologists do not withhold the data on which their conclusions are based from other competent professionals who seek to verify the substantive claims through reanalysis and who intend to use such data only for that purpose, provided that the confidentiality of the participants can be protected and unless legal rights concerning proprietary data preclude their release" (Standard 8.14).

APA expects authors to adhere to these standards. Specifically, APA expects authors to have their data available throughout the editorial review process and for at least 5 years after the date of publication.

Authors are required to state in writing that they have complied with APA ethical standards in the treatment of their sample, human or animal, or to describe the details of treatment.

Download Certification of Compliance With APA Ethical Principles Form (PDF, 26KB)


Other Information

Appeals Process for Manuscript Submissions

Preparing Auxiliary Files for Production

Document Deposit Procedures for APA Journals