DOI:
10.1016/j.neubiorev.2017.04.010

Document Version
Peer reviewed version

Title: Meta-analysis of the effects of intranasal oxytocin on interpretation and expression of emotions

Authors: Jenni Leppanen, Kah Wee Ng, Kate Tchanturia, Janet Treasure

PII: S0149-7634(16)30744-8
DOI: http://dx.doi.org/doi:10.1016/j.neubiorev.2017.04.010
Reference: NBR 2820

To appear in:

Received date: 22-11-2016
Revised date: 10-4-2017
Accepted date: 11-4-2017

Please cite this article as: Leppanen, Jenni, Ng, Kah Wee, Tchanturia, Kate, Treasure, Janet, Meta-analysis of the effects of intranasal oxytocin on interpretation and expression of emotions.Neuroscience and Biobehavioral Reviews http://dx.doi.org/10.1016/j.neubiorev.2017.04.010

Meta-analysis of the effects of intranasal oxytocin on interpretation and expression of emotions

Ms. Jenni Leppanen[1]*, Dr. Kah Wee Ng[2], Dr. Kate Tchanturia[1,3], Prof. Janet Treasure[1]

[1] King's College London, Institute of Psychiatry, Psychology, and Neuroscience, Department of Psychological Medicine

[2] Singapore General Hospital, 20 College Road, Academia, Singapore, 169865

[3] Department of Psychology, Illia State University, Tbilisi, Georgia

* Corresponding author:

Jenni Leppanen, Department of Psychological Medicine, Institute of Psychiatry, Psychology, & Neuroscience, King's College London, 103 Denmark Hill, London, SE5 8AF, UK. Email: jenni.leppanen@kcl.ac.uk

Highlights

- Single dose of intranasal oxytocin has little to no effect on emotional theory of mind or emotion expression among healthy or clinical adult populations
- Intranasal oxytocin improved recognition of basic emotions, particularly fear, but only among healthy populations
- Intranasal oxytocin improves sensitivity to recognise anger, but only among healthy populations
- Effect of oxytocin may be dependent on social boundaries

1 Abstract

Accurate interpretation and appropriate expression of emotions are key aspects of social-cognition. Several mental disorders are characterised by transdiagnostic difficulties in these areas and, recently, there has been increasing interest in exploring the effects of oxytocin on social-emotional functioning.

This review consists of 33 studies. Fifteen of the studies included people with autism spectrum disorder, schizophrenia, borderline personality disorder, frontotemporal dementia, anorexia nervosa, bulimia nervosa, post-traumatic stress disorder, depression, and opioid and alcohol dependence. We conducted ten meta-analyses examining the effects of intranasal oxytocin on expression of emotions, emotional theory of mind, sensitivity to recognise basic emotions, and recognition of basic emotions.

A single dose of intranasal oxytocin significantly improved the recognition of basic emotions, particularly fear, and increased the expression of positive emotions among the healthy individuals. Oxytocin did not significantly influence theory of mind or the expression of negative emotions among the healthy individuals. Finally, intranasal oxytocin did not significantly influence interpretation or expression of emotions among the clinical populations.

Keywords: oxytocin, theory of mind, emotion recognition, emotion expression

2 Introduction

Accurate interpretation of other's emotions, appropriate expression of one's own emotions, and reciprocity within interactions are key aspects of social cognition. In social interaction, emotion expression is dependent on accurate interpretation of social signals (Hess and Fischer, 2013; Künecke et al., 2014). According to the embodied simulation theory, emotion expression and mimicry, in turn, play an important role in facilitating the interpretation of others' expressions, empathy, and prosocial behaviour in recipients (Gallese, 2005). Indeed, behavioural studies have documented that automatic mimicry of emotions facilitates recognition, whereas blocking mimicry impairs recognition accuracy and sensitivity (Argaud et al., 2016; Duffy and Chartrand, 2015; Künecke et al., 2014; Rychlowska et al., 2014; Schneider et al., 2013). Anomalies in emotion expression also have social and affective consequences, with incongruent emotion expression increasing the desire for greater social distance and negative social evaluation by the recipient (Brown et al., 2015; Szczurek et al., 2012). Similarly, expressive suppression has been found to increase the suppressors' blood pressure, subjective anxiety, and social isolation (Butler et al., 2003; Gross, 2002).

Anomalies in social-emotional functioning are important transdiagnostic features in several psychiatric disorders (Bora and Berk, 2016; Bora and Köse, 2016; Chung et al., 2014; Davies et al., 2016; Henry et al., 2014; Kring and Moran, 2008). Meta-analyses have found that people with eating disorders (EDs), depression, schizophrenia, and autism spectrum disorders (ASD) have similar difficulties in accurate interpretation of emotions, including recognition of basic emotions in faces and tone of voice with small

effect sizes and in emotional theory of mind with medium to large effect sizes (Bora and Berk, 2016; Bora and Köse, 2016; Caglar-Nazali et al., 2014; Chung et al., 2014; Uljarevic and Hamilton, 2013). Recent systematic reviews have also found that people with schizophrenia, EDs, depression, ASD, and borderline personality disorder (PBD) display less positive facial affect in response to positive emotional stimuli (Davies et al., 2016; Kring and Moran, 2008). Furthermore, a recent meta-analysis of 537 task-based fMRI studies in depression, bipolar disorder, schizophrenia, and obsessive-compulsive disorder failed to find significant differences between the disorders in whole brain neural response to social and cognitive tasks (Sprooten et al., 2016). Together these findings suggest that anomalies in social-emotional processing in psychiatric disorders may have shared underlying mechanisms. Given social and affective consequences of these difficulties, better understanding of the underlying processes is of interest. One such possible mechanism is the oxytocin system.

Preclinical studies have found that the neuropeptide, oxytocin, may regulate social-emotional functioning (Dölen et al., 2013; Hicks et al., 2012; Lim and Young, 2006; Lukas et al., 2011; Onaka et al., 2012). Endogenous oxytocin has been found to play an important role in the central and medial amygdala, facilitating formation of social bonds, maternal behaviour, and social recognition in rodents (Lim and Young, 2006; Onaka et al., 2012). Additionally, a recent study found that formation of social reward was dependent on coordinated activity between oxytocin and serotonin in the mouse nucleus accumbens (Dölen et al., 2013). In rodents, the administration of oxytocin receptor agonist and exogenous synthetic oxytocin has also been found to increase social place preference and reduce social defeat induced avoidance (Hicks et al., 2012;

Lukas et al., 2011). Conversely, the administration of oxytocin receptor antagonist, has been found to increase corticosteroid levels and induce social avoidance in monkeys during times of stress (Cavanaugh et al., 2016).

Recently there has been increasing interest in translating these findings into humans and the effects of intranasal oxytocin on social-emotional function has been widely studied (Bakermans-Kranenburg and van Ijzendoorn, 2013; Bartz et al., 2011; Guastella and MacLeod, 2012; Shahrestani et al., 2013; van Ijzendoorn and Bakermans-Kranenburg, 2012). A few previous meta-analytic reviews have found that intranasal oxytocin improves recognition of anger and happiness, and increases in-group trust among healthy individuals with small effect sizes (Shahrestani et al., 2013; van Ijzendoorn and Bakermans-Kranenburg, 2012). However, to our knowledge no meta-analyses to date have investigated the effects of a single dose of oxytocin on recognition of all six basic emotions, other aspects of emotion interpretation, including theory of mind or sensitivity to recognise basic emotions, or on emotion expression among healthy individuals.

To date, two meta-analytic reviews have investigated the effects of intranasal oxytocin on different aspects of social-emotional functioning in a variety of clinical groups (Bakermans-Kranenburg and van Ijzendoorn, 2013; Ooi et al., 2017). One reported small, but generally positive effect of intranasal oxytocin on social-emotional functioning and psychopathology among people with ASD, anxiety disorders, depression, schizophrenia, and BPD (Bakermans-Kranenburg and van Ijzendoorn, 2013; Ooi et al., 2017). The other meta-analysis found no significant effects of

intranasal oxytocin on social-emotional processing in ASD (Ooi et al., 2017). However, these reviews were quite heterogeneous pooling studies assessing psychopathology and social-emotional processing, or single dose and repeated dose studies into one meta-analysis (Bakermans-Kranenburg and van Ijzendoorn, 2013; Ooi et al., 2017). To our knowledge no previous meta-analyses have investigated the effects of a single dose of intranasal oxytocin separately on different aspects of interpretation and expression of emotions among both healthy and clinical populations. In order to consider the possibility of translating animal studies more widely into treatment for psychiatric disorders it is important to consider various key outcomes and whether there is evidence that they might be modified by oxytocin.

The aim of the current review was to pool studies investigating the effects of a single dose of intranasal oxytocin on various aspects of social-emotional functioning among healthy and clinical populations. Specifically, we aimed to examine the effects of intranasal oxytocin on theory of mind, recognition of basic emotions, sensitivity to recognise basic emotions, and on emotion expression among healthy and clinical populations. We tested the hypothesis that oxytocin would improve all aspects of social-emotional functioning.

3 Methodology

3.1 Literature searches

Electronic databases, including OVID (journals@OVID, PsycINFO, PsycARTICLES, Embase, AGRIS, MEDLINE), PubMed, and Web of Knowledge core collection, were searched for studies published during available years up to February 2017 in accordance with the PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). Two separate literature searches were conducted in order to uncover studies investigating the effects of a single dose of intranasal oxytocin on interpretation and expression of emotions in a social context. The first literature search was conducted with the following search terms: *oxytocin AND emotion AND (interpretation OR recognition OR identification OR labelling OR "theory of mind" OR mentalising OR perception OR empath*)*. The second search was conducted with the following search terms: *oxytocin AND emotion AND (expression OR mimicry OR mirroring OR communication OR responsiveness OR expressivity)*. Additionally, to ensure no studies were missed by the initial search, the bibliographies of included studies were searched for additional studies.

3.2 Eligibility criteria

Studies were included if they met the following inclusion criteria: 1) investigated the effects of a single dose of intranasal exogenous oxytocin on interpretation or expression of emotions in a social context among either healthy adult participants or adult clinical populations (18 years old or older); 2) compared the effects of intranasal oxytocin with intranasal placebo spray; 3) investigated short term outcomes; and 4) randomly allocated participants to oxytocin and placebo groups or, in the case of

crossover, within subjects studies, randomised the treatment order. Any studies, which used tasks that did not include a social component, such as the bumper car theory of mind task where social context is inferred from the movement of triangles on a computer screen, were excluded. Trials, in which participants either received repeated doses of oxytocin or in which long term outcomes of a single dose of oxytocin were assessed, were excluded. Studies that included only children or adolescents were excluded, because the majority of them were longer trials and the effects of oxytocin on social-emotional processing can be different in adults and children. Full-text articles published in peer reviewed journals and where possible, published conference abstracts were included.

In total, five studies were excluded after further screening because they incorporated tasks that were very different compared to the other included studies despite being otherwise relevant. These studies included a theory of mind task involving infant stimuli, continuously assessing the mood of a target on a video clip on a 9-point Likert scale, manipulating the context in which the emotions were presented, recognising emotions from a point-light-display, and interpreting basic emotions from tone of voice in different languages (Bartz et al., 2010; Bernaerts et al., 2016; De Dreu et al., 2016; Perry et al., 2013; Voorthuis et al., 2014).

3.3 Study selection

The literature searches were conducted by one author (J.L.). The studies yielded from the literature search were then screened based on their titles and abstracts. Full text

articles were then assessed for eligibility followed by final screening and assessment by two authors (J.L. and K.W.N.). Where appropriate conference abstracts of studies not yet published were also screened and assessed for eligibility. If deemed eligible the authors were contacted in order to gain access to the data. Only studies that both authors agreed on were included in the final systematic review and meta-analyses. Any cases where eligibility remained in question were brought to the whole team for further discussion and assessment. The study selection processes of the two searches are presented in Figures 1 and 2.

Figure 1. PRISMA flow chart of selection of articles for interpretation of emotions

----------FIGURE 1----------

Figure 2. PRISMA flow chart of selection of articles for expression of emotions

----------FIGURE 2----------

3.4 Data collection and synthesis

Seventeen studies reported their results in figures or otherwise did not include the relevant data in the paper or supplementary materials. The corresponding authors of these papers were contacted by K.W.N. in order to gain access to the relevant data. The authors of the following papers provided the required data via personal correspondence: Aoki et al. (2014), Averbeck et al. (2012), Cardoso et al. (2014a), Chen et al. (2015), Fischer-Shofty et al. (2010), Kirkpatrick et al. (2014), Koch et al. (2016), Korb et al. (2016), and Luminet et al. (2011).

In order to conduct meta-analyses means, standard deviations, and samples sizes for both oxytocin and placebo groups or sessions were extracted from the studies or acquired through personal correspondence. Where standard error of the mean was reported, standard deviation was estimated with the following formula $SD = SE * \sqrt{N}$. Altogether ten meta-analyses were conducted to investigate the effects of a single dose of intranasal oxytocin on interpretation and expression of emotions. The first meta-analysis investigated the effects of intranasal oxytocin on emotional theory of mind among healthy and clinical populations. The second meta-analysis investigated the effects of a single dose of intranasal oxytocin on interpretation of basic emotions among healthy and clinical populations. Five out of fifteen studies included in the overall basic emotion recognition meta-analysis included all six basic emotions. Thus, to provide further information about the effects of intranasal oxytocin on each of the six basic emotions, separate meta-analyses were conducted on recognition of anger, fear, disgust, sadness, surprise, and happiness among healthy and clinical populations. The last two meta-analyses investigated the effects of intranasal oxytocin on emotion recognition sensitivity and expression of congruent emotions in response to emotionally provoking stimuli. The clinical category included people with ASD, schizophrenia, depression, anorexia nervosa (AN), bulimia nervosa (BN), BPD, behavioural variant of frontotemporal dementia (FTD), alcohol dependence disorder, and opioid dependence disorder.

Additional information regarding age, the oxytocin dose (in international units [IU]), the specific diagnosis of clinical participants, the proportion of female participants in

the sample, and type of task used were also recorded. We also included information regarding the presence of ceiling effects in accuracy scores. Since there are currently no guidelines to indicate what should be used as a cut-off for ceiling effects, we chose an arbitrary cut-off of 85%. Studies, in which the accuracy scores were greater than or equal to 85%, were coded as having evidence of ceiling effects. The additional information was used to conduct meta-regressions to identify variables that might explain any potential between-study heterogeneity.

3.5 Emotion interpretation tasks

The emotion interpretation tasks included are summarised in Table 1. The theory of mind tasks included the Reading the Mind in the Eyes task (RMET) (Baron-Cohen et al., 2001), Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) (Mayer et al., 2003), The Awareness of Social Inference Test (TASIT) (McDonald et al., 2003), and the Sally Anne task (Baron-Cohen et al., 1985). In the RMET participants are presented with 36 photographs of eyes and asked to select one of four words that best describes the complex emotion displayed. The MSCEIT has six subscales, three of which were of interest for the current review and assessed perception of emotions in others and understanding of how emotions change and blend together. The TASIT has three subscales, two of which were of interest for the current review and assessed emotional evaluation and social-emotional inference based on verbal and visual stimuli. Finally, the Sally Anne task assesses emotional and cognitive theory of mind with black and white comic strips. The current reviewed only included data on emotional theory of mind.

The studies that investigated the effects of a single dose of intranasal oxytocin on interpretation of basic emotions mostly used standard emotion recognition paradigms in which participants were presented with a photograph of a face depicting a basic emotion and asked to identify the emotion. A subset of studies used dynamic emotion recognition tasks, in which participants were presented with a neutral face that gradually morphed into the full emotion, or masked emotion recognition tasks, in which the photograph of a face depicting a basic emotion was either preceded or followed by a neutral face. The outcome was recognition accuracy. All studies used standardised stimuli to assess emotion recognition accuracy.

All studies examining interpretation sensitivity used similar dynamic emotion recognition tasks as described above. In these tasks participants were initially presented with a neutral face that gradually morphed into full emotion. The outcome was estimated as the intensity percentage at which participants accurately recognised the emotion. Lower the intensity percentage, the more sensitive the participants were to detect the emotion. All studies used standardised stimuli to assess emotion recognition sensitivity.

3.6 Emotion expression tasks

The emotion expression tasks included are summarised in Table 1. All included studies used different paradigms to elicit positive and negative emotions. One of the studies presented participants with emotionally provoking positive and negative images from the International Affective Picture System (IAPS). Two of the studies presented participants with short film clips. In one study the film clips depicted either a happy

face that gradually morphed into an angry expression or an angry face that gradually morphed into a happy expression. The other study presented participants with excerpts from movies that were designed to elicit happy and sad emotions. Lastly, one study assessed expressions of flight and affiliation during a clinical interview. Participants facial expression were recorded either by manually assessing facial expression using the Facial Expression Coding System (FACES, (Kring and Sloan, 1991)), using facial electromyography (EMG), or using automated facial emotion detection software, Noldus FaceReader (Noldus Information Technology b.v., www.noldus.com).

3.7 Statistical analysis

All statistical analysis was performed using R (R Core Team, 2015). For studies using between subjects design Hedges' g was calculated to estimate unbiased effect size with 95% confidence intervals. For studies using within subjects design standardised mean change (SMC) with change score standardisation was calculated to estimate effect size with 95% confidence intervals. The SMC controls for the correlation in task performance between the two assessments (oxytocin session and placebo session). Where correlation between the two assessments was not reported, the correlation coefficient was estimated using the following formula $r = \frac{SD1^2+SD2^2-SDchange^2}{2\times SD1\times SD2}$ (Morris and DeShon, 2002). Both Hedges' g and SMC effect size estimates was are on the same scale and were interpreted as small (≥0.20), medium (≥0.50), and large (≥0.80) (Hedges, 1981). In nine of the ten meta-analyses higher scores indicated either greater accuracy or greater facial expressivity. Thus, in these meta-analyses positive effect size indicated improved emotion interpretation accuracy or increased emotion

expression following oxytocin administration, whereas negative effect size indicated poorer emotion interpretation or reduced emotion expression following oxytocin administration. In the meta-analysis investigating emotion recognition sensitivity lower scores indicated greater sensitivity to recognise the facial expressions. Thus, in this meta-analysis negative effect sizes indicated increased sensitivity to recognise emotions following oxytocin administration and positive effect sizes indicated reduced sensitivity to recognise emotions following oxytocin administration. Significance threshold was set at $p < 0.05$ unless otherwise stated.

The user contributed Metafor package was used to conduct the meta-analyses, meta-regressions, and publication bias estimation (Viechtbauer, 2010). The meta-analyses were conducted with a multivariate random effects model with an autoregressive structure using the *rma.mv* function in Metafor to account for correlations arising from multiple outcomes from the same sample. Between-study heterogeneity was assessed by calculating *Cochran's Q* index and I^2 index. Where between-study heterogeneity was found meta-regressions were conducted and the moderator effects of the following variables was assessed: age, the dose administered (in IU), specific diagnosis of clinical participants, the type of task used, the proportion of female participants in the sample, and whether there were ceiling effects present (accuracy ≥ 85%). All binary and categorical moderators were dummy coded and entered into the meta-regression using the *factor* function. The impact of each moderator was assessed in separate models.

Influential studies and extreme outliers were identified by inspecting Cook's distance plots and the standardised residuals of each individual study. Where the z-score of the standardised residuals exceeded 1.96, the study was deemed to be an outlier (Viechtbauer and Cheung, 2010). To test the impact of any outliers on the pooled effect size estimate, we conducted the meta-analysis with the outlier present followed by a meta-regression to examine whether the outlier significantly explained the between study heterogeneity. If the outlier significantly explained the heterogeneity, the outlier was excluded and only results from the meta-analysis without the outlier were reported in full. This procedure led to exclusion of one study (Xu et al., 2015) from two separate meta-analyses.

Publication bias was investigated with Begg's rank correlation test for funnel plot asymmetry (Begg and Mazumdar, 1994) and where significant effects were present Rosenthal's file drawer analysis was also conducted (Rosenthal, 1979). The robustness of the significant findings was assessed by calculating the Rosenthal's criterion ($5n+10$, n = the number of studies in the meta-analysis) and comparing that figure to the fail-safe N from the file drawer analysis. If the fail-safe N exceeded the criterion the findings were considered robust, if it did not this was taken as an indicator of publication bias.

Finally, the statistical power of each study included in the review was assessed by comparing the sample sizes in these studies against that recommended by a power calculation conducted with G*Power (Faul et al., 2007). According to the power calculation studies using between subjects design should have at least 64 participants

in each group while studies using within subjects design should have at least 34 participants altogether, to have adequate statistical power (≥80%) to reliably detect a moderate difference (ES ≥ 0.5) between the two groups or conditions.

4 Results

4.1 Study characteristics

The characteristics of the 33 included studies are summarised in Table 1. The effect size estimate (ES) represents standardised mean difference (Hedges' g) or standardised mean change (SMC) in interpretation and expression of emotions following oxytocin and placebo. The last column in Table 1 indicates whether the study met the sample size requirement for adequate statistical power (≥80%) to reliably detect at least a medium sized effect between the oxytocin and placebo groups or conditions. None of the studies that used between subjects design met the requirement for adequate power, but five of the studies that used within subjects design did meet this requirement.

Table 1. Study characteristics

----------TABLE 1----------

4.2 Effects of oxytocin on interpretation of emotions

4.2.1 Theory of mind

Fourteen studies were included in the meta-analysis investigating the effects of a single dose of intranasal oxytocin on emotional theory of mind (Figure 3). Six of the

studies included clinical populations, namely individuals with ASD, AN, depression, schizophrenia, and FTD, and two studies included individuals with substance dependence disorder, including opioid and alcohol dependence. Overall, there was no significant effect of oxytocin on theory of mind (ES = 0.09, Z = 1.14, p = 0.256, 95% CI [-0.06, 0.24], k = 23, N of levels = 17). When the healthy and clinical groups were inspected separately there was no evidence of significant effect of intranasal oxytocin on emotional theory of mind within either group (Healthy: ES = 0.07, Z = 0.76, p = 0.447, 95% CI [-0.10, 0.24], k = 13, N of levels = 9; Clinical: ES = 0.10, Z = 0.74, p = 0.457, 95% CI [-0.17, 0.37], k = 10, N of levels = 8).

There was evidence of significant between-study heterogeneity (Q = 65.85, p < 0.0001, I^2 = 58.42%), which was further explored with meta-regressions. The findings from the meta-regressions were corrected for multiple comparisons with Bonferroni correction (0.05/6) and p < 0.008 was considered significant. The meta-regressions revealed a significant moderator effect of age (Qm = 7.33, p = 0.007; Supplementary Figure 1), still leaving some residual heterogeneity (Qr = 34.26, p = 0.012). This finding suggests that younger participants showed greater oxytocin-induced improvement in emotional theory of mind. The dose administered (Qm = 1.04, p = 0.309; Supplementary Figure 2), the type of task used (Qm = 2.78, p = 0.427), the diagnostic group (Qm = 0.01, p = 0.918), and the proportion of female participants in the sample (Qm = 1.36, p = 0.243; Supplementary Figure 3) did not significantly explain the between-study heterogeneity.

Begg's rank correlation test of funnel plot asymmetry approached significance suggesting there may have been some publication bias (T = 0.29, p = 0.057; Supplementary Figure 4).

Figure 3. Effects of oxytocin on theory of mind

----------FIGURE 3----------

4.2.2 Recognition of basic emotions

Seventeen studies investigated the effects of intranasal oxytocin on total basic emotion recognition accuracy. Based on standardised residuals and Cook's distance, one study by Xu et al. (2015) was identified as an extreme outlier (standardised res. = 2.68, Z = 4.56, SE = 0.59; Supplementary Table 1; Supplementary Figure 5). The impact of the outlier on the meta-analysis was investigated by conducting a meta-analysis with the outlier present, which yielded a significant oxytocin-induced improvement in recognition of basic emotions with a small effect size along with significant between-study heterogeneity (ES = 0.28, Z = 2.56, p = 0.011, 95% CI [0.07, 0.50], k = 33, N of levels = 23, Q = 99.98, p < 0.0001, I^2 = 81.12%). We then investigated whether the outlier was significantly different from the other studies by conducting a meta-regression, which revealed that the outlier significantly explained the between-study heterogeneity (Qm = 46.98, p < 0.0001) leaving no significant residual heterogeneity (Qr = 43.51, p = 0.067). Therefore, the outlier was removed from further analysis.

Full meta-analysis was conducted with the remaining sixteen studies (Figure 4). Five of the studies included clinical populations, namely people with schizophrenia, BPD,

ASD, and post-traumatic stress disorder. The meta-analysis revealed that intranasal oxytocin administration improved overall basic emotion interpretation accuracy with a negligible effect size (ES = 0.18, Z = 3.09, p = 0.002, 95% CI [0.06, 0.29], k = 32, N of levels = 22). When the effects of intranasal oxytocin on basic emotions recognition were further investigated within the healthy and clinical populations, the results showed that oxytocin significantly improved basic emotion recognition among the healthy individuals with a negligible effect size (ES = 0.13, Z = 2.30, p = 0.022, 95% CI [0.02, 0.24], k = 23, N of levels = 17). Among the mixed clinical population oxytocin-induced improvement on basic emotion recognition approached significance with a small effect size (ES = 0.27, Z = 1.82, p = 0.069, 95% CI [-0.02, 0.57], k = 9, N of levels = 5).

The between-study heterogeneity approached significance (Q = 42.08, p = 0.089, I^2 = 30.27%), and was therefore, explored further with meta-regressions. The findings from the meta-regressions were corrected for multiple comparisons with Bonferroni correction (0.05/7) and $p < 0.007$ was considered significant. The proportion of female participants in the sample (Qm = 1.34, p = 0.247; Supplementary Figure 6), the diagnostic group (Qm = 3.47, p = 0.482), age (Qm = 0.67, p = 0.414; Supplementary Figure 7), the dose administered (Qm = 0.49, p = 0.486; Supplementary Figure 8), the type of task used (Qm = 1.32, p = 0.858), and the presence of ceiling effects (Qm = 1.48, p = 0.224) did not significantly explain the between-study heterogeneity.

Begg's rank correlation test of funnel plot asymmetry did not reveal significant publication bias (T = -0.08, p = 0.506; Supplementary Figure 9). Rosenthal's the file

drawer analysis indicated a fail-safe N of 182, suggesting that 182 studies finding no significant effects of oxytocin on recognition of basic emotions would be required to reduce the observed effects to null. This exceeds Rosenthal's criterion for this meta-analysis (5*n*+10=90) suggesting that the effect was quite robust.

Since the total scores in most of the included studies did not consist of all six basic emotions further meta-analyses were conducted to examine if he above effect was driven by a subset of basic emotions. Thus, we conducted six additional meta-analyses to investigate the effects of a single dose of intranasal oxytocin on the recognition of anger, fear, disgust, sadness, surprise, and happiness. Where possible the studies were divided into healthy and clinical subgroups, which were both inspected separately.

Figure 4. Effect of oxytocin on overall basic emotion recognition

----------FIGURE 4----------

4.2.3 Recognition of anger

Ten studies investigated the effects of intranasal oxytocin on recognition of anger (Figure 5). Two of these studies included people with schizophrenia and BPD.

Overall, oxytocin administration did not significantly improve recognition of anger (ES = 0.05, Z = 1.03, p = 0.305, 95% CI [-0.05, 0.15], k = 18, N of levels = 11). When the groups were inspected separately, there was no significant effect of oxytocin on

recognition of anger within the healthy (ES = 0.04, Z = 0.71, p = 0.476, 95% CI [-0.07, 0.14], k = 14, N of levels = 9) or clinical groups (ES = 0.17, Z = 1.10, p = 0.275, 95% CI [-0.13, 0.47], k = 4, N of levels = 2).

There was no evidence of significant between-study heterogeneity (Q = 21.16, p = 0.219, I^2 = 3.45e-09%) There was also no evidence of significant publication bias on the Begg's rank correlation test of funnel plot asymmetry (T = 0.03, p = 0.881; Supplementary Figure 10).

Figure 5. Effect of oxytocin on recognition of anger.

----------FIGURE 5----------

4.2.4 Recognition of fear

Nine studies investigated the effects of intranasal oxytocin on recognition of fear (Figure 6). Two of the studies included people with schizophrenia and BPD.

Overall, oxytocin administration significantly improved recognition of fear with a small effect size (ES = 0.21, Z = 2.95, p = 0.003, 95% CI [0.07, 0.34], k = 14, N of levels = 10). This effect was driven by a significant oxytocin-induced improvement in recognition of fear among the healthy individuals with a small effect size (ES = 0.24, Z = 2.63, p = 0.009, 95% CI [0.06, 0.41], k = 10, N of levels = 8). There was no significant effect of

oxytocin among the mixed clinical population (ES = 0.16, Z = 1.05, p = 0.295, 95% CI [-0.14, 0.46]).

The between-study heterogeneity approached significance (Q = 19.95, p = 0.096, I^2 = 20.92%) and was therefore explored further with meta-regressions. The findings from the meta-regressions were corrected for multiple comparisons with Bonferroni correction (0.05/7) and p < 0.007 was considered significant. The dose administered (Qm = 4.19, p = 0.041; Supplementary Figure 11), the proportion of female participants in the sample (Qm = 0.23, p = 0.632; Supplementary Figure 12), the presence of ceiling effects (Qm = 0.05, p = 0.831), age (Qm = 2.56, p = 0.109; Supplementary Figure 13), the type of task used (Qm = 0.78, p = 0.676), and the diagnostic group (Qm = 0.90, p = 0.638) did not significantly explain the between-study heterogeneity.

There was no evidence of significant publication bias on the Begg's rank correlation test of funnel plot asymmetry (T = -0.08, p = 0.747; Supplementary Figure 14). Rosenthal's file drawer analysis revealed a fail-safe N of 59, suggesting that 59 additional studies reporting no effect of oxytocin would be needed to reduce the observed effect to null. This exceeds Rosenthal's criterion for this meta-analysis (5*n*+10 = 55), suggesting the effect was quite robust.

Figure 6. Effect of oxytocin on recognition of fear.

----------FIGURE 6----------

4.2.5 Recognition of disgust

Four studies investigated the effects of oxytocin on the recognition of disgust (Figure 7). Because only one of these studies included clinical populations, namely people with schizophrenia, it was not possible to investigate the effects of oxytocin separately within the healthy and clinical populations. Thus, the data was analysed across groups.

The meta-analysis showed a negligible oxytocin-induced improvement in the recognition of disgust, which approached significance (ES = 0.18, Z = 1.73, p = 0.083, 95% CI [-0.02, 0.39], k = 5, N of levels = 4).

There was no evidence of significant between-study heterogeneity (Q = 2.07, p = 0.722, I^2 = 2.73e-08%).

There was no evidence of significant publication bias on the Begg's rank correlation test of funnel plot asymmetry (T = 0.20, p = 0.817; Supplementary Figure 15).

Figure 7. Effect of oxytocin on recognition of disgust.

----------FIGURE 7----------

4.2.6 Recognition of sadness

Seven studies investigated the effects of oxytocin on the recognition of sadness (Figure 8). As above, since only one study included a clinical group, people with schizophrenia, the meta-analysis was conducted across groups.

The meta-analysis revealed no significant effects of oxytocin on the recognition of sadness (ES = 0.04, Z = 0.53, p = 0.594, 95% CI [-0.10, 0.17]). There was also no significant between-study heterogeneity (Q = 7.59, p = 0.474, I^2 = 16.67%).

There was no evidence of significant publication bias on the Begg's rank correlation test of funnel plot asymmetry (T = 0.33, p = 0.260; Supplementary Figure 16).

Figure 8. Effect of oxytocin on recognition of sadness.

----------FIGURE 8----------

4.2.7 Recognition of surprise

Four studies investigated the effects of oxytocin on the recognition of surprise (Figure 9). Since only one study included a clinical group, people with schizophrenia, the meta-analysis was conducted across groups.

The meta-analysis revealed no significant effect of oxytocin on the recognition of surprise (ES = -0.02, Z = -0.16, p = 0.874, 95% CI [-0.22, 0.19], k = 5, N of levels = 4). There was also no evidence of significant between-study heterogeneity (Q = 3.72, p = 0.445, I^2 = 1.90e-08%)

There was no evidence of significant publication bias on Begg's rank correlation test of funnel plot asymmetry (T = 0.20, p = 0.817; Supplementary Figure 17).

Figure 9. Effect of oxytocin on recognition of surprise.

----------FIGURE 9----------

4.2.8 Recognition of happiness

Twelve studies investigated the effects on intranasal oxytocin on the recognition of happiness. Following inspection of the standardised residuals and Cook's distance one study by Xu et al. (2015) was identified as an extreme outlier (standardised res. = 2.83, Z = 3.27, SE = 0.87; Supplementary Table 2, Supplementary Figure 18). The meta-analysis with the outlier present yielded no significant effect of oxytocin on the recognition of happiness, but there was significant between-study heterogeneity (ES = 0.31, Z = 1.26, p = 0.207, 95% CI [-0.17, 0.79], k = 20, N of levels = 12, Q = 93.93, p < 0.0001, I^2 = 93.34%). A meta-regression revealed that the outlier significantly explained the between-study heterogeneity (Qm = 51.44, p < 0.0001) leaving no significant residual heterogeneity (Qr = 28.36, p = 0.057). Therefore, the outlier was removed from further analysis.

The eleven remaining studies were included in the final meta-analysis (Figure 10). Two of the studies included clinical groups, people with schizophrenia and BPD. Overall, oxytocin did not significantly improve the recognition of happiness (ES = 0.08, Z = 1.07, p = 0.284, 95% CI [-0.07, 0.23], k = 19, N of levels = 12). When the healthy and clinical groups were inspected separately, there were no differential effects among the healthy (ES = 0.10, Z = 1.18, p = 0.237, 95% CI [-0.06, 0.26], k = 15, N of levels = 10), or mixed clinical populations (ES = -0.07, Z = -0.23, p = 0.821, 95% CI [-0.72, 0.57]).

The between-study heterogeneity approached significance (Q = 27.87, p = 0.064, I^2 = 33.62%) and was explored further with meta-regressions. The findings from the meta-regressions were corrected for multiple comparisons with Bonferroni correction (0.05/7) and p < 0.007 was considered significant. The type of task used (Qm = 11.36, p = 0.010), the proportion of female participants in the sample (Qm = 0.84, p = 0.360; Supplementary Figure 19), age (Qm = 0.07, p = 0.789; Supplementary Figure 20), the dose administered (Qm = 2.67, p = 0.103; Supplementary Figure 21), the diagnostic group (Qm = 2.84, p = 0.242), and the presence of ceiling effects (Qm = 3.29, p = 0.070), did not significantly explain the between-study heterogeneity.

There was no evidence of significant publication bias on the Begg's rank correlation test of funnel plot asymmetry (T = -0.13, p = 0.447; Supplementary Figure 22).

Figure 10. Effect of oxytocin on recognition of happiness.

----------FIGURE 10----------

4.2.9 Emotion recognition sensitivity:

Three studies investigated the effects of intranasal oxytocin on emotion recognition sensitivity (Figure 11). Because only one of these studies included clinical groups, people with AN and BN, the meta-analysis was conducted across groups.

Overall, intranasal oxytocin did not significantly influence emotion recognition sensitivity (ES = -0.14, Z = -1.45, p = 0.146, 95% CI [-0.34, 0.05], k = 20, N of levels = 5).

We then investigated recognition sensitivity of each emotion separately and found that oxytocin did not significantly improve the sensitivity to recognise happiness (ES = -0.11, Z = -1.24, p = 0.216, 95% CI [-0.29, 0.07], k = 5, N of levels = 5), sadness (ES = -0.17, Z = -1.25, p = 0.212, 95% CI [-0.43, 0.09], k = 5, N of levels = 5), fear (ES = -0.03, Z = -0.42, p = 0.674, 95% CI [-0.20, 0.13], k = 5, N of levels = 5), or anger (ES = -0.20, Z = -1.64, p = 0.100, 95% CI [-0.43, 0.04], k = 5, N of levels = 5).

The between-study heterogeneity approached significance (Q = 29.31, p = 0.061, I^2 = 67.31%) and was explored further with meta-regressions. The findings from the meta-regressions were corrected for multiple comparisons with Bonferroni correction (0.05/5) and $p < 0.01$ was considered significant. The proportion of female participants in the sample (Qm = 0.03, p = 0.864; Supplementary Figure 23), age (Qm = 1.27, p = 0.261; Supplementary Figure 24), the dose administered (Qm = 0.03, p = 0.864; Supplementary Figure 25), and the diagnostic group (Qm = 2.03, p = 0.567) did not significantly explain the heterogeneity.

Begg's rank correlation test of funnel plot asymmetry revealed evidence of significant publication bias (T = -0.74, $p < 0.0001$; Supplementary Figure 26) indicating that small studies finding large effects were more likely to be published.

Figure 11. Effect of oxytocin on emotion recognition sensitivity.

----------FIGURE 11----------

4.3 Effects of oxytocin on emotion expression

Four studies investigated the effects of a single dose of intranasal oxytocin on expression of congruent emotions in response to emotionally provoking stimuli (Figure 12). Three studies included individuals with clinical conditions, including BPD, Schizophrenia, and AN.

Overall, there was no significant effect of oxytocin on emotion expression (ES = 0.08, Z = 0.87, p = 0.385, 95% CI [-0.10, 0.26], k = 14, N of levels = 7). When the data was further inspected, the meta-analysis showed that oxytocin significantly increased the expression of positive emotions among the healthy individuals with a small effect size (ES = 0.25, Z = 2.29, p = 0.022, 95% CI [0.04, 0.47], k = 4, N of levels = 4). Oxytocin did not significantly influence the expression of negative emotions among the healthy individuals (ES = 0.10, Z = 0.55, p = 0.585, 95% CI [-0.27, 0.47], k = 4, N of levels = 4). There was also no evidence of significant effects of oxytocin on the expression of positive or negative emotions among the clinical populations (Positive ES = 0.02, Z = 0.13, p = 0.896, 95% CI [-0.22, 0.25], k = 3, N of levels = 3; Negative: ES = -0.08, Z -0.35, p = 0.726, 95% CI [-0.52, 0.36], k = 3, N of levels = 3).

There was significant between-study heterogeneity (Q = 23.14, p = 0.040, I^2 = 35.95%), which was explored further with meta-regressions. The proportion of female participants in the sample, age, and the dose administered individually significantly explained the heterogeneity (Supplementary Figures 27-29). These variables were, thus, entered into a full model which significantly explained the between-study heterogeneity (Qm = 8.80, p = 0.032) leaving no significant residual heterogeneity (Qr

= 14.33, p = 0.158). This finding suggests that older male participants who received 24IU of intranasal oxytocin showed greater oxytocin induced increase in facial expressivity. However, this meta-analysis consisted of only four studies and, thus, this finding should be interpreted with caution. The diagnostic group (Qm = 6.14, p = 0.105)did not significantly explain the heterogeneity.

There was no evidence of significant publication bias on the Begg's rank correlation test of funnel plot asymmetry (T = -0.03, p = 0.915; Supplementary Figure 30). However, Rosenthal's file drawer analysis revealed a fail-safe N of 1, indicating that only 1 study finding no significant effects of oxytocin on emotion expression would be needed to reduce the observed significant effect to null. This does not exceed the Rosenthal's criterion for this meta-analysis (5*k*+10 = 30) indicating that this finding is not robust and there is likely to be substantial publication bias present.

Figure 12. Effect of oxytocin on emotion expression.

----------FIGURE 12----------

5 Discussion

The aim of the current meta-analytic review was to investigate the effects of a single dose of intranasal oxytocin on interpretation and expression of emotions among healthy and clinical populations. Most the studies recruited only healthy individuals, but thirteen studies also included people with clinical disorders and two studies included people with substance dependence disorder. The meta-analyses revealed

that a single dose of intranasal oxytocin significantly improved the recognition of basic emotions, particularly fear, but only among healthy individuals with small to negligible effect sizes. Oxytocin also increased the expression of positive emotions with a small effect size among the healthy individuals. Intranasal oxytocin did not significantly influence theory of mind among the healthy individuals. Although, the oxytocin-induced improvement in basic emotion recognition approached significance, overall there were no significant effects on intranasal oxytocin on the interpretation or expression of emotions among the mixed clinical population.

The oxytocin-induced increase in the expression of positive emotions in healthy individuals is in line with findings from previous systematic reviews and suggests that oxytocin plays an important role in facilitating prosocial behaviour in humans (Bakermans-Kranenburg and van Ijzendoorn, 2013; Churchland and Winkielman, 2012; Guastella and MacLeod, 2012). This finding is also supported by previous work that has documented increased trust and cooperation particularly towards the members of a safe “in-group” following oxytocin administration (van Ijzendoorn and Bakermans-Kranenburg, 2012). Together these findings suggest that oxytocin may facilitate pro-social behaviour among healthy individuals. However, it is of importance to note that the oxytocin-induced increase in the expression of positive emotions was not robust, and it is likely that there was substantial publication bias present.

The present findings also showed that intranasal oxytocin improved recognition of basic emotions, which appeared to be largely driven by oxytocin-induced significant improvement in fear recognition and nearly significant improvement in disgust

recognition. These results may seem surprising in this light of the findings above. However, these results are in line with a previous meta-analysis, which also found that intranasal oxytocin improved early recognition of angry faces and late recognition of fearful faces among healthy individuals (Shahrestani et al., 2013). Additionally, intranasal oxytocin has also been found to increase attention and approach towards negative social-emotional stimuli, such as angry and fearful facial expressions (Clark-Elford et al., 2015; Simon et al., 2012; Tollenaar et al., 2013).

Recent systematic reviews have attempted to explain the recent emergence of similar seemingly contradictory findings by suggesting that the effects of intranasal oxytocin may be modulated by social boundaries (Olff et al., 2013; Zik and Roberts, 2015). This hypothesis suggests that oxytocin may be of evolutionary importance in social interactions, increasing pro-social behaviour towards "safe" stimuli and defensiveness towards "unsafe" stimuli (Olff et al., 2013; Zik and Roberts, 2015). Thus, when faced with positive social-emotional cues oxytocin may facilitate pro-social behaviour, trust, and cooperation. Conversely, when presented with unfamiliar or negative stimuli oxytocin administration may increase attention and alertness towards these potentially threatening, negative social-emotional cues. Although this hypothesis has not been previously linked to the effects of oxytocin on social-emotional processing, it is consistent with the present findings and is supported by several behavioural studies using investment and trust games. These studies have documented that intranasal oxytocin increases empathy towards "in-group" members as well as cooperation and compliance within the "in-group" (De Dreu and Kret, 2016; Ten Velden et al., 2017). Oxytocin also reduced cooperation with "out-group" members

even when the “out-group” members are generous towards the participants (Daughters et al., 2017).

Although, the oxytocin-induced improvement in basic emotion recognition approached significance, overall the present series of meta-analyses found no significant oxytocin-induced changes in interpretation or expression of emotions among the mixed clinical population. One possible explanation is that the different disorders included in the mixed clinical group were too heterogeneous and that there were too few studies with the same disorder group to draw firm conclusions. Indeed, there was significant heterogeneity between the diagnostic groups in the present series of meta-analyses. However, there were also substantial differences in effect size estimates between studies that included the same patient group, suggesting that there may be individual differences within the diagnostic groups that modulate the effects of oxytocin.

Large-scale cohort studies have documented that clinical populations have inter-individual variability with substantial heterogeneity in types of comorbidity and aetiological risk (Lamers et al., 2010; Melartin et al., 2002; Sterling et al., 2008; Wessman et al., 2009). A few recent systematic reviews have also suggested that the effects of oxytocin on social-emotional functioning may be moderated by contextual and individual differences (Bartz et al., 2011; Olff et al., 2013). Factors such as attachment style and experience of early parental care have been found to moderate the effects of oxytocin; those scoring low on attachment avoidance and harsh parenting showed greater oxytocin-induced increase in social cooperation and

positive response to crying infants (Bakermans-Kranenburg et al., 2011; Fang et al., 2014; Olff et al., 2013). Thus, further investigation of potential contextual and individual differences that may modulate the effects of oxytocin is of interest.

5.1 Limitations, recommendations and future directions

The majority of the studies included in the present review did not meet the sample size requirement for adequate statistical power to reliably detect at least a moderate effect of the drug. Although meta-analyses are a statistically powerful method to pool studies and increase statistical power (Greco et al., 2013), it is important that individual studies also have adequate power. This has been suggested to be a particularly big problem in oxytocin research and the potential source of the wide range of different and sometimes contradictory findings arising from different studies (Walum et al., 2016). These problems may go some way to explain the between-study heterogeneity and publication bias in the present review. Therefore, we recommend that future studies should recruit much larger number of participants to ensure that reliable effects of intranasal oxytocin can be detected.

Several studies investigating the effects of a single dose of intranasal oxytocin on recognition of basic emotions had evidence of ceiling effects with mean accuracy percentages over 85%. Even though presence of ceiling effects did not have significant impact on the present findings, such effects make it difficult to find the true effect of the drug or intervention because performance is already at maximum. Future studies investigating the effects of intranasal oxytocin on the recognition of basic emotions,

should opt for alternative tasks, for example presenting images where the emotion is less than 100% present.

Seven of the meta-analyses in the present review had evidence of between-study heterogeneity and in three of them we were unable to identify the source of the heterogeneity. Previous studies have reported several confounding factors such as individual differences in anxiety and attachment security that can influence the effects of oxytocin on social-emotional processing (Bakermans-Kranenburg et al., 2011; Fang et al., 2014; Olff et al., 2013). The impact of these factors were not explored or reported in majority of the studies included in the present review meaning that it was not possible for us to explore the impact of these factors had on the meta-analyses. We recommend that future studies explore the impact of individual differences to gain better understanding of the impact of potential confounding factors on the effects of intranasal oxytocin.

There has also recently been some doubt regarding the effects of intranasal oxytocin on social-emotional functioning in general and criticism directed at the lack of compelling theoretical framework to explain the contradictory findings reported thus far (Lane et al., 2016; Leng and Ludwig, 2016). Some important questions have also been raised regarding if and how intranasal oxytocin accesses the brain and in what quantities (Leng and Ludwig, 2016). These criticisms have become increasingly important to address in the light of recent replication failures (Lane et al., 2015; Radke and de Bruijn, 2015). We have taken some steps in the present review to introduce a potential theoretical framework, but several issues still remain unanswered. For

instance, there are a number of methodological obstacles, such as individual differences in the nasal cavity physiology, the nasal spray formulations used, and the devices used to deliver the drug, that need to be considered (Quintana et al., 2016). Furthermore, there are uncertainties regarding the site and mechanism of action of exogenous intranasal oxytocin in humans due to lack of suitable radio tracers (Paloyelis et al., 2016). There has also been some suggestion that oxytocin has important effects in the periphery, moderating cardiovascular functioning and the peripheral cortisol response, which may influence social behaviour (Cardoso et al., 2014b; Gutkowska et al., 2014). Further research is needed to answer these important questions particularly if oxytocin is to be used as a treatment enhancer to support clinical care.

Finally, no correction for multiple comparisons was applied in the present series of meta-analyses. Although we took steps to combat this issue by investigating the robustness of all significant findings and advising caution when interpreting the moderator analyses, the presence of false positive findings cannot be ruled out.

6 Conclusions

The current meta-analytic review pooled studies investigating the effects of a single dose of intranasal oxytocin on social-cognition. There was no significant effect of intranasal oxytocin on interpretation or expression of emotions among the mixed clinical population. Intranasal oxytocin significantly improved recognition of basic emotions, particularly fear, and increased the expression of positive emotions, but

only among the healthy individuals. These findings are in line with previous work that has found that the effects of oxytocin may be modulated by social boundaries. Further large-scale research is needed to better understand of the role of intranasal oxytocin in social-emotional processing and potential moderators of its' effects.

Acknowledgements

JL is supported by a scholarship from the Psychiatry Research Trust. KWN is supported by Singapore General Hospital. KT received funding from Swiss Anorexia Foundation. JT receives salary support from the National Institute for Health Research (NIHR), Mental Health Biomedical Research at South London and Maudsley NHS Foundation Trust, and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

The authors do not report any financial or non-financial conflicting interests.

7 References

Aoki, Y., Yahata, N., Watanabe, T., Takano, Y., Kawakubo, Y., Kuwabara, H., Iwashiro, N., Natsubori, T., Inoue, H., Suga, M., Takao, H., Sasaki, H., Gonoi, W., Kunimatsu, A., Kasai, K., Yamasue, H., 2014. Oxytocin improves behavioural and neural deficits in inferring others' social emotions in autism. Brain : a journal of neurology 137, 3073-3086.

Argaud, S., Delplanque, S., Houvenaghel, J.-F., Auffret, M., Duprez, J., Vérin, M., Grandjean, D., Sauleau, P., 2016. Does Facial Amimia Impact the Recognition of Facial Emotions? An EMG Study in Parkinson's Disease. PloS one 11, e0160329.

Averbeck, B.B., Bobin, T., Evans, S., Shergill, S.S., 2012. Emotion recognition and oxytocin in patients with schizophrenia. Psychological medicine 42, 259-266.

Bakermans-Kranenburg, M.J., van Ijzendoorn, M.H., 2013. Sniffing around oxytocin: review and meta-analyses of trials in healthy and clinical groups with implications for pharmacotherapy. Translational psychiatry 3, e258.

Bakermans-Kranenburg, M.J., van IJzendoorn, M.H., Riem, M.M.E., Tops, M., Alink, L.R.A., 2011. Oxytocin decreases handgrip force in reaction to infant crying in females without harsh parenting experiences. Social cognitive and affective neuroscience.

Baron-Cohen, S., Leslie, A.M., Frith, U., 1985. Does the autistic child have a "theory of mind" ? Cognition 21, 37-46.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I., 2001. The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. Journal of child psychology and psychiatry, and allied disciplines 42, 241-251.

Bartz, J.A., Zaki, J., Bolger, N., Hollander, E., Ludwig, N.N., Kolevzon, A., Ochsner, K.N., 2010. Oxytocin selectively improves empathic accuracy. Psychological science 21, 1426-1428.

Bartz, J.A., Zaki, J., Bolger, N., Ochsner, K.N., 2011. Social effects of oxytocin in humans: context and person matter. Trends in cognitive sciences 15, 301-309.

Begg, C.B., Mazumdar, M., 1994. Operating characteristics of a rank correlation test for publication bias. Biometrics, 1088-1101.

Bernaerts, S., Berra, E., Wenderoth, N., Alaerts, K., 2016. Influence of oxytocin on emotion recognition from body language: A randomized placebo-controlled trial. Psychoneuroendocrinology 72, 182-189.

Bora, E., Berk, M., 2016. Theory of mind in major depressive disorder: A meta-analysis. Journal of Affective Disorders 191, 49-55.

Bora, E., Köse, S., 2016. Meta-analysis of theory of mind in anorexia nervosa and bulimia nervosa: A specific İmpairment of cognitive perspective taking in anorexia nervosa? International Journal of Eating Disorders 49, 739-740.

Brown, C.M., Olkhov, Y.M., Bailey, V.S., Daniels, E.R., 2015. Gender Differences in the Social Cost of Affective Deviance. The Journal of social psychology 155, 535-540.

Butler, E.A., Egloff, B., Wlhelm, F.H., Smith, N.C., Erickson, E.A., Gross, J.J., 2003. The social consequences of expressive suppression. Emotion 3, 48.

Caglar-Nazali, H.P., Corfield, F., Cardi, V., Ambwani, S., Leppanen, J., Olabintan, O., Deriziotis, S., Hadjimichalis, A., Scognamiglio, P., Eshkevari, E., Micali, N., Treasure, J., 2014. A systematic review and meta-analysis of 'Systems for Social Processes' in eating disorders. Neuroscience and biobehavioral reviews 42, 55-92.

Cardoso, C., Ellenbogen, M.A., Linnen, A.M., 2014a. The Effect of Intranasal Oxytocin on Perceiving and Understanding Emotion on the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). Emotion 14, 43-50.

Cardoso, C., Kingdon, D., Ellenbogen, M.A., 2014b. A meta-analytic review of the impact of intranasal oxytocin administration on cortisol concentrations during laboratory tasks: moderation by method and mental health. Psychoneuroendocrinology 49, 161-170.

Cavanaugh, J., Carp, S.B., Rock, C.M., French, J.A., 2016. Oxytocin modulates behavioral and physiological responses to a stressor in marmoset monkeys. Psychoneuroendocrinology 66, 22-30.

Chen, F.S., Kumsta, R., Dvorak, F., Domes, G., Yim, O.S., Ebstein, R.P., Heinrichs, M., 2015. Genetic modulation of oxytocin sensitivity: a pharmacogenetic approach. Translational psychiatry 5, e664.

Chung, Y.S., Barch, D., Strube, M., 2014. A Meta-Analysis of Mentalizing Impairments in Adults With Schizophrenia and Autism Spectrum Disorder. Schizophrenia bulletin 40, 602-616.

Churchland, P.S., Winkielman, P., 2012. Modulating social behavior with oxytocin: How does it work? What does it mean? Hormones and behavior 61, 392-399.

Clark-Elford, R., Nathan, P.J., Auyeung, B., Mogg, K., Bradley, B.P., Sule, A., Muller, U., Dudas, R.B., Sahakian, B.J., Baron-Cohen, S., 2015. Effects of oxytocin on attention to emotional faces in healthy volunteers and highly socially anxious males. The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP) 18.

Daughters, K., Manstead, A.S.R., Ten Velden, F.S., De Dreu, C.K.W., 2017. Oxytocin modulates third-party sanctioning of selfish and generous behavior within and between groups. Psychoneuroendocrinology 77, 18-24.

Davies, H., Wolz, I., Leppanen, J., Fernandez-Aranda, F., Schmidt, U., Tchanturia, K., 2016. Facial expression to emotional stimuli in non-psychotic disorders: A systematic review and meta-analysis. Neuroscience and biobehavioral reviews 64, 252-271.

De Dreu, C.K.W., Kret, M.E., 2016. Oxytocin Conditions Intergroup Relations Through Upregulated In-Group Empathy, Cooperation, Conformity, and Defense. Biological psychiatry 79, 165-173.

De Dreu, C.K.W., Kret, M.E., Sauter, D.A., 2016. Assessing Emotional Vocalizations From Cultural In-Group and Out-Group Depends on Oxytocin. Social Psychological and Personality Science 7, 837-846.

Dölen, G., Darvishzadeh, A., Huang, K.W., Malenka, R.C., 2013. Social reward requires coordinated activity of accumbens oxytocin and 5HT. Nature 501, 179-184.

Duffy, K.A., Chartrand, T.L., 2015. Mimicry: causes and consequences. Current Opinion in Behavioral Sciences 3, 112-116.

Fang, A., Hoge, E.A., Heinrichs, M., Hofmann, S.G., 2014. Attachment Style Moderates the Effects of Oxytocin on Social Behaviors and Cognitions During Social Rejection: Applying a Research Domain Criteria Framework to Social Anxiety. Clinical Psychological Science 2, 740-747.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods 39, 175-191.

Fischer-Shofty, M., Shamay-Tsoory, S.G., Harari, H., Levkovitz, Y., 2010. The effect of intranasal administration of oxytocin on fear recognition. Neuropsychologia 48, 179-184.

Gallese, V., 2005. Embodied simulation: From neurons to phenomenal experience. Phenomenology and the Cognitive Sciences 4, 23-48.

Greco, T., Zangrillo, A., Biondi-Zoccai, G., Landoni, G., 2013. Meta-analysis: pitfalls and hints. Heart, Lung and Vessels 5, 219-225.

Gross, J.J., 2002. Emotion regulation: Affective, cognitive, and social consequences. Psychophysiology 39, 281-291.

Guastella, A.J., MacLeod, C., 2012. A critical review of the influence of oxytocin nasal spray on social cognition in humans: evidence and future directions. Hormones and behavior 61, 410-418.

Gutkowska, J., Jankowski, M., Antunes-Rodrigues, J., 2014. The role of oxytocin in cardiovascular regulation. Brazilian Journal of Medical and Biological Research 47, 206-214.

Hedges, L.V., 1981. Distribution Theory for Glass's Estimator of Effect size and Related Estimators. Journal of Educational and Behavioral Statistics 6, 107-128.

Henry, J.D., Phillips, L.H., von Hippel, C., 2014. A meta-analytic review of theory of mind difficulties in behavioural-variant frontotemporal dementia. Neuropsychologia 56, 53-62.

Hess, U., Fischer, A., 2013. Emotional mimicry as social regulation. Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc 17, 142-157.

Hicks, C., Jorgensen, W., Brown, C., Fardell, J., Koehbach, J., Gruber, C.W., Kassiou, M., Hunt, G.E., McGregor, I.S., 2012. The Nonpeptide Oxytocin Receptor Agonist WAY 267,464: Receptor-Binding Profile, Prosocial Effects and Distribution of c-Fos Expression in Adolescent Rats. Journal of neuroendocrinology 24, 1012-1029.

Kirkpatrick, M.G., Lee, R., Wardle, M.C., Jacob, S., de Wit, H., 2014. Effects of MDMA and Intranasal oxytocin on social and emotional processing. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 39, 1654-1663.

Koch, S.B.J., van Zuiden, M., Nawijn, L., Frijling, J.L., Veltman, D.J., Olff, M., 2016. Intranasal Oxytocin Administration Dampens Amygdala Reactivity towards Emotional Faces in Male and Female PTSD Patients. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 41, 1495-1504.

Korb, S., Malsert, J., Strathearn, L., Vuilleumier, P., Niedenthal, P., 2016. Sniff and mimic - Intranasal oxytocin increases facial mimicry in a sample of men. Hormones and behavior 84, 64-74.

Kring, A.M., Moran, E.K., 2008. Emotional response deficits in schizophrenia: insights from affective science. Schizophrenia bulletin 34, 819-834.

Kring, A.M., Sloan, D., 1991. The facial expression coding system (FACES): A users guide. Unpublished manuscript.

Künecke, J., Hildebrandt, A., Recio, G., Sommer, W., Wilhelm, O., 2014. Facial EMG Responses to Emotional Expressions Are Related to Emotion Perception Ability. PloS one 9, e84053.

Lamers, F., de Jonge, P., Nolen, W.A., Smit, J.H., Zitman, F.G., Beekman, A.T., Penninx, B.W., 2010. Identifying depressive subtypes in a large cohort study: results from the

Netherlands Study of Depression and Anxiety (NESDA). The Journal of clinical psychiatry 71, 1582-1589.

Lane, A., Luminet, O., Nave, G., Mikolajczak, M., 2016. Is there a Publication Bias in Behavioural Intranasal Oxytocin Research on Humans? Opening the File Drawer of One Laboratory. Journal of neuroendocrinology 28.

Lane, A., Mikolajczak, M., Treinen, E., Samson, D., Corneille, O., de Timary, P., Luminet, O., 2015. Failed Replication of Oxytocin Effects on Trust: The Envelope Task Case. PloS one 10, e0137000.

Leng, G., Ludwig, M., 2016. Intranasal Oxytocin: Myths and Delusions. Biological psychiatry 79, 243-250.

Lim, M.M., Young, L.J., 2006. Neuropeptidergic regulation of affiliative behavior and social bonding in animals. Hormones and behavior 50, 506-517.

Lukas, M., Toth, I., Reber, S.O., Slattery, D.A., Veenema, A.H., Neumann, I.D., 2011. The Neuropeptide Oxytocin Facilitates Pro-Social Behavior and Prevents Social Avoidance in Rats and Mice. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 36, 2159-2168.

Luminet, O., Grynberg, D., Ruzette, N., Mikolajczak, M., 2011. Personality-dependent effects of oxytocin: greater social benefits for high alexithymia scorers. Biological psychology 87, 401-406.

Mayer, J.D., Salovey, P., Caruso, D.R., Sitarenios, G., 2003. Measuring emotional intelligence with the MSCEIT V2.0. Emotion 3, 97-105.

McDonald, S., Flanagan, S., Rollins, J., Kinch, J., 2003. TASIT: A new clinical tool for assessing social perception after traumatic brain injury. The Journal of head trauma rehabilitation 18, 219-238.

Melartin, T.K., Rytsala, H.J., Leskela, U.S., Lestela-Mielonen, P.S., Sokero, T.P., Isometsa, E.T., 2002. Current comorbidity of psychiatric disorders among DSM-IV major depressive disorder patients in psychiatric care in the Vantaa Depression Study. The Journal of clinical psychiatry 63, 126-134.

Morris, S.B., DeShon, R.P., 2002. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological methods 7, 105-125.

Olff, M., Frijling, J.L., Kubzansky, L.D., Bradley, B., Ellenbogen, M.A., Cardoso, C., Bartz, J.A., Yee, J.R., van Zuiden, M., 2013. The role of oxytocin in social bonding, stress regulation and mental health: An update on the moderating effects of context and interindividual differences. Psychoneuroendocrinology 38, 1883-1894.

Onaka, T., Takayanagi, Y., Yoshida, M., 2012. Roles of Oxytocin Neurones in the Control of Stress, Energy Metabolism, and Social Behaviour. Journal of Neuroendocrinology 24, 587-598.

Ooi, Y.P., Weng, S.-J., Kossowsky, J., Gerger, H., Sung, M., 2017. Oxytocin and Autism Spectrum Disorders: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Pharmacopsychiatry 50, 5-13.

Paloyelis, Y., Doyle, O.M., Zelaya, F.O., Maltezos, S., Williams, S.C., Fotopoulou, A., Howard, M.A., 2016. A Spatiotemporal Profile of In Vivo Cerebral Blood Flow Changes Following Intranasal Oxytocin in Humans. Biological psychiatry 79, 693-705.

Perry, A., Aviezer, H., Goldstein, P., Palgi, S., Klein, E., Shamay-Tsoory, S.G., 2013. Face or body? Oxytocin improves perception of emotions from facial expressions in incongruent emotional body context. Psychoneuroendocrinology 38, 2820-2825.

Quintana, D.S., Guastella, A.J., Westlye, L.T., Andreassen, O.A., 2016. The promise and pitfalls of intranasally administering psychopharmacological agents for the treatment of psychiatric disorders. Mol Psychiatry 21, 29-38.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Radke, S., de Bruijn, E.R., 2015. Does oxytocin affect mind-reading? A replication study. Psychoneuroendocrinology 60, 75-81.

Rosenthal, R., 1979. The file drawer problem and tolerance for null results. Psychological Bulletin 86, 638-641.

Rychlowska, M., Cañadas, E., Wood, A., Krumhuber, E.G., Fischer, A., Niedenthal, P.M., 2014. Blocking Mimicry Makes True and False Smiles Look the Same. PloS one 9, e90876.

Schneider, K.G., Hempel, R.J., Lynch, T.R., 2013. That “poker face” just might lose you the game! The impact of expressive suppression and mimicry on sensitivity to facial expressions of emotion. Emotion 13, 852-866.

Shahrestani, S., Kemp, A.H., Guastella, A.J., 2013. The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: A meta-analysis. Neuropsychopharmacology 38, 1929-1936.

Simon, N.M., Lynn, S., Hoge, E., Fischer, L., Barrett, L.F., 2012. Oxytocin influences response bias in men but not women in a signal detection emotion perception task. Neuropsychopharmacology 38, S357-S358.

Sprooten, E., Rasgon, A., Goodman, M., Carlin, A., Leibu, E., Lee, W.H., Frangou, S., 2016. Addressing reverse inference in psychiatric neuroimaging: Meta-analyses of

task-related brain activation in common mental disorders. Human Brain Mapping, n/a-n/a.

Sterling, L., Dawson, G., Estes, A., Greenson, J., 2008. Characteristics Associated with Presence of Depressive Symptoms in Adults with Autism Spectrum Disorder. Journal of autism and developmental disorders 38, 1011-1018.

Szczurek, L., Monin, B., Gross, J.J., 2012. The Stranger effect: the rejection of affective deviants. Psychol Sci 23, 1105-1111.

Ten Velden, F.S., Daughters, K., De Dreu, C.K.W., 2017. Oxytocin promotes intuitive rather than deliberated cooperation with the in-group. Hormones and behavior.

Tollenaar, M.S., Chatzimanoli, M., van der Wee, N.J., Putman, P., 2013. Enhanced orienting of attention in response to emotional gaze cues after oxytocin administration in healthy young men. Psychoneuroendocrinology 38, 1797-1802.

Uljarevic, M., Hamilton, A., 2013. Recognition of Emotions in Autism: A Formal Meta-Analysis. Journal of autism and developmental disorders 43, 1517-1526.

van Ijzendoorn, M.H., Bakermans-Kranenburg, M.J., 2012. A sniff of trust: Meta-analysis of the effects of intranasal oxytocin administration on face recognition, trust to in-group, and trust to out-group. Psychoneuroendocrinology 37, 438-443.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J Stat Softw 36, 1-48.

Viechtbauer, W., Cheung, M.W.L., 2010. Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods 1, 112-125.

Voorthuis, A., Riem, M.M., Van Ijzendoorn, M.H., Bakermans-Kranenburg, M.J., 2014. Reading the mind in the infant eyes: Paradoxical effects of oxytocin on neural activity

and emotion recognition in watching pictures of infant faces. Brain research 1580, 151-159.

Walum, H., Waldman, I.D., Young, L.J., 2016. Statistical and Methodological Considerations for the Interpretation of Intranasal Oxytocin Studies. Biological psychiatry 79, 251-257.

Wessman, J., Paunio, T., Tuulio-Henriksson, A., Koivisto, M., Partonen, T., Suvisaari, J., Turunen, J.A., Wedenoja, J., Hennah, W., Pietiläinen, O.P.H., Lönnqvist, J., Mannila, H., Peltonen, L., 2009. Mixture Model Clustering of Phenotype Features Reveals Evidence for Association of DTNBP1 to a Specific Subtype of Schizophrenia. Biological psychiatry 66, 990-996.

Xu, L., Ma, X., Zhao, W., Luo, L., Yao, S., Kendrick, K.M., 2015. Oxytocin enhances attentional bias for neutral and positive expression faces in individuals with higher autistic traits. Psychoneuroendocrinology 62, 352-358.

Zik, J.B., Roberts, D.L., 2015. The many faces of oxytocin: implications for psychiatry. Psychiatry research 226, 31-37.

8 Figure legends

Figure 1. PRISMA flow chart of the study selection process for the emotion interpretation meta-analyses.

Figure 2. PRISMA flow chart of the study selection process for the emotion expression meta-analyses.

Figure 3. Effects of intranasal oxytocin vs. placebo on theory of mind. Positive effect sizes indicate improved theory of mind following oxytocin administration; negative effect sizes indicate reduced theory of mind following oxytocin administration. RMET = Reading the mind in the eyes test; MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; TASIT = The Awareness of Social Inference Test; AN = anorexia nervosa; ASD = autism spectrum disorder; frontotemporal dementia = FTD.

Figure 4. Effect of oxytocin vs. placebo on overall basic emotion recognition. Positive effect sizes indicate improved emotion recognition following oxytocin administration; negative effect sizes indicate reduced emotion recognition following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces 30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity; DANVA = Diagnostic Analysis of Non-Verbal Accuracy; ASD = Autism spectrum disorder; BPD = Borderline personality disorder; PTSD = Post-traumatic stress disorder. Kirkpatrick, et al. 2014: a = 20IU of intranasal oxytocin, b = 40IU of intranasal oxytocin. Campbell, et al. 2014a = young female participants, Campbell, et al. 2014b = young male participants, Campbell, et al. 2014c = older female participants, Campbell, et al. 2014d = older male participants.

Figure 5. Effect of oxytocin vs. placebo on recognition of anger. Positive effect sizes indicate improved recognition of anger following oxytocin administration; negative effect sizes indicate reduced recognition of anger following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces 30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity; BPD = Borderline personality disorder. Kirkpatrick, et al. 2014: a = 20IU of intranasal oxytocin, b = 40IU of intranasal oxytocin.

Figure 6. Effects of oxytocin vs. placebo on recognition of fear. Positive effect sizes indicate improved recognition of fear following oxytocin administration; negative effect sizes indicate reduced recognition of fear following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces 30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity.

Figure 7. Effect of oxytocin vs. placebo on recognition of disgust. Positive effect sizes indicate improved recognition of disgust following oxytocin administration; negative effect sizes indicate reduced recognition of disgust following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces 30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity; BPD = Borderline personality disorder. Kirkpatrick, et al. 2014: a = 20IU of intranasal oxytocin, b = 40IU of intranasal oxytocin.

Figure 8. Effect of oxytocin vs. placebo on recognition of sadness. Positive effect sizes indicate improved recognition of sadness following oxytocin administration; negative effect sizes indicate reduced recognition of sadness following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces 30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity. Kirkpatrick, et al. 2014: a = 20IU of intranasal oxytocin, b = 40IU of intranasal oxytocin.

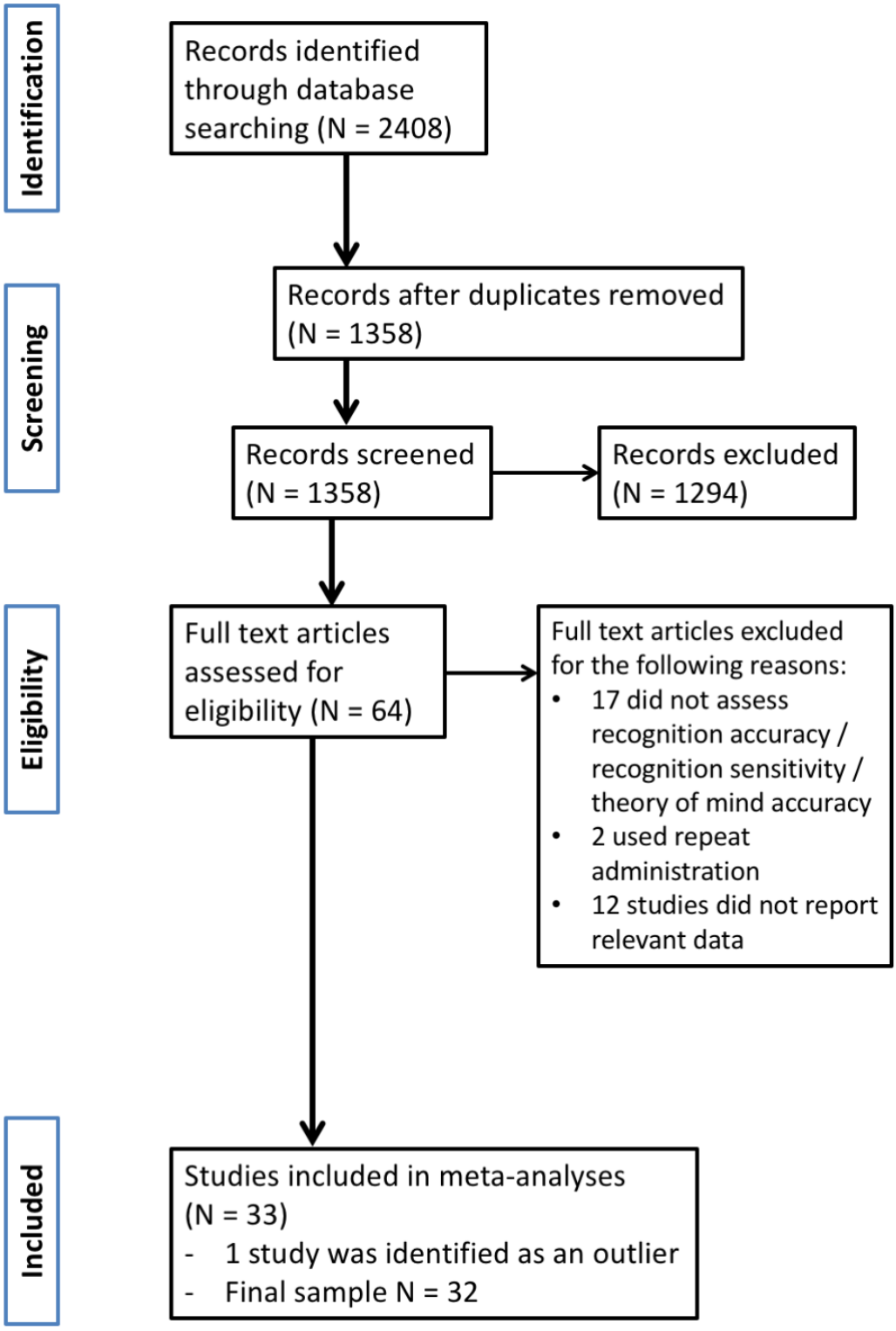
Figure 9. Effect of oxytocin vs. placebo on recognition of surprise. Positive effect sizes indicate improved recognition of surprise following oxytocin administration; negative effect sizes indicate reduced recognition of surprise following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces

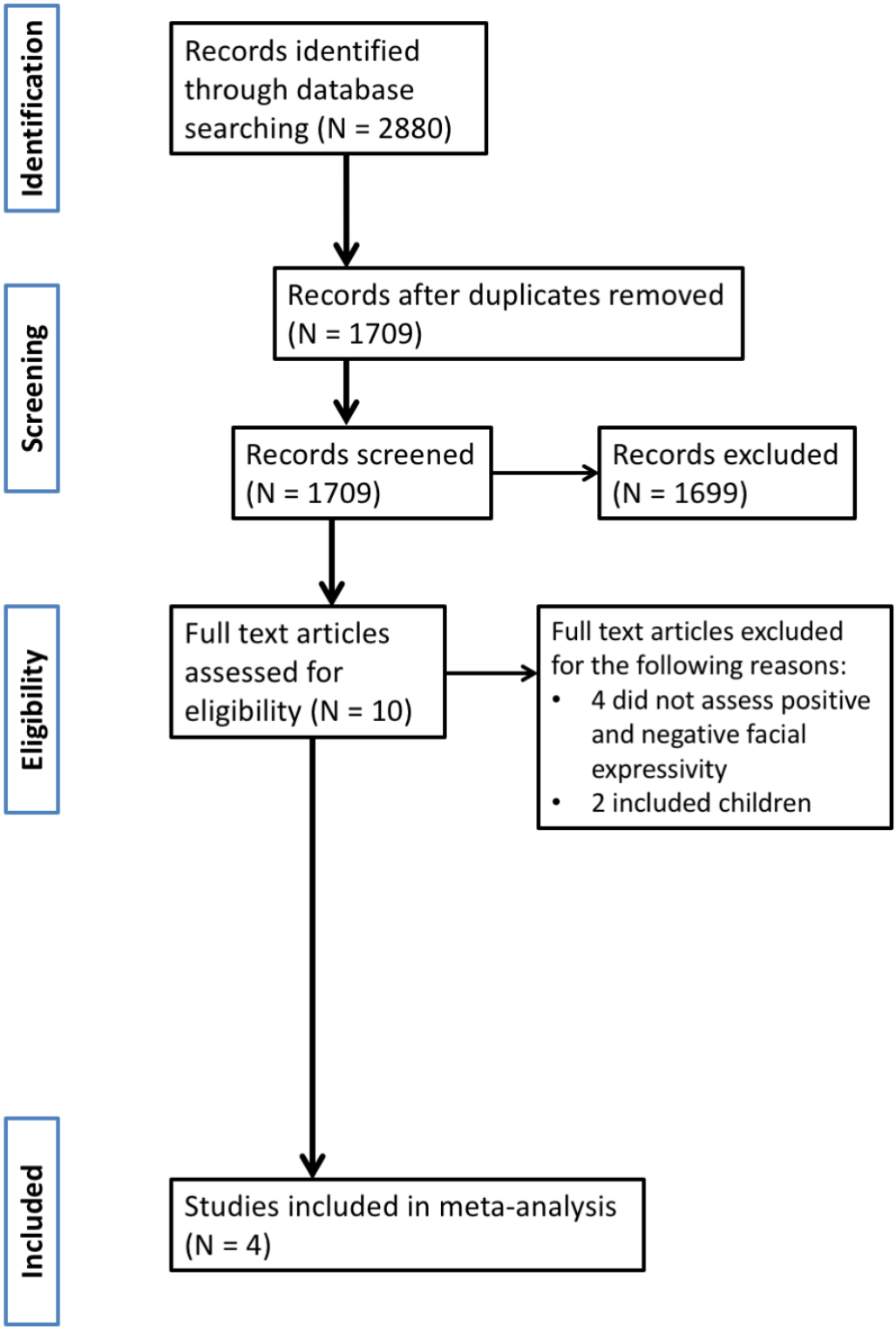
30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity.

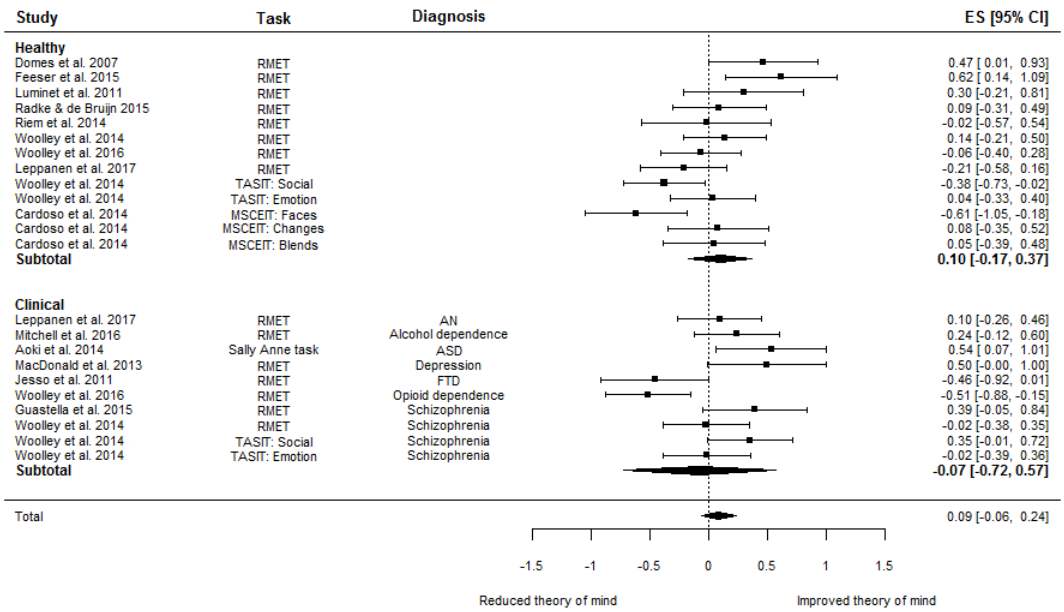
Figure 10. Effect of oxytocin vs. placebo on recognition of happiness. Positive effect sizes indicate improved recognition of happiness following oxytocin administration; negative effect sizes indicate reduced recognition of happiness following oxytocin administration. HEDT: 30% = Hexagon emotion discrimination task: morphed faces 30% intensity; HEDT: 70% = Hexagon emotion discrimination task: morphed faces 70% intensity; BPD = Borderline personality disorder. Kirkpatrick, et al. 2014: a = 20IU of intranasal oxytocin, b = 40IU of intranasal oxytocin.

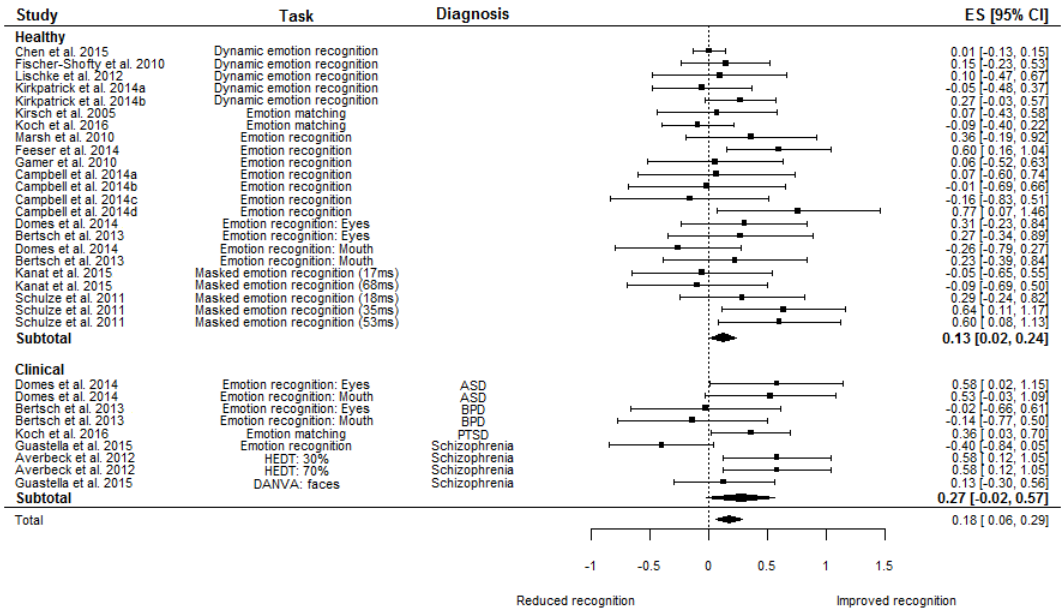
Figure 11. Effect of oxytocin vs. placebo on emotion recognition sensitivity. Negative effect sizes indicate improved emotion recognition sensitivity following oxytocin administration; positive effect sizes indicate reduced emotion recognition sensitivity following oxytocin administration. All studies used dynamic emotion recognition task. BN = Bulimia nervosa, AN = Anorexia nervosa.

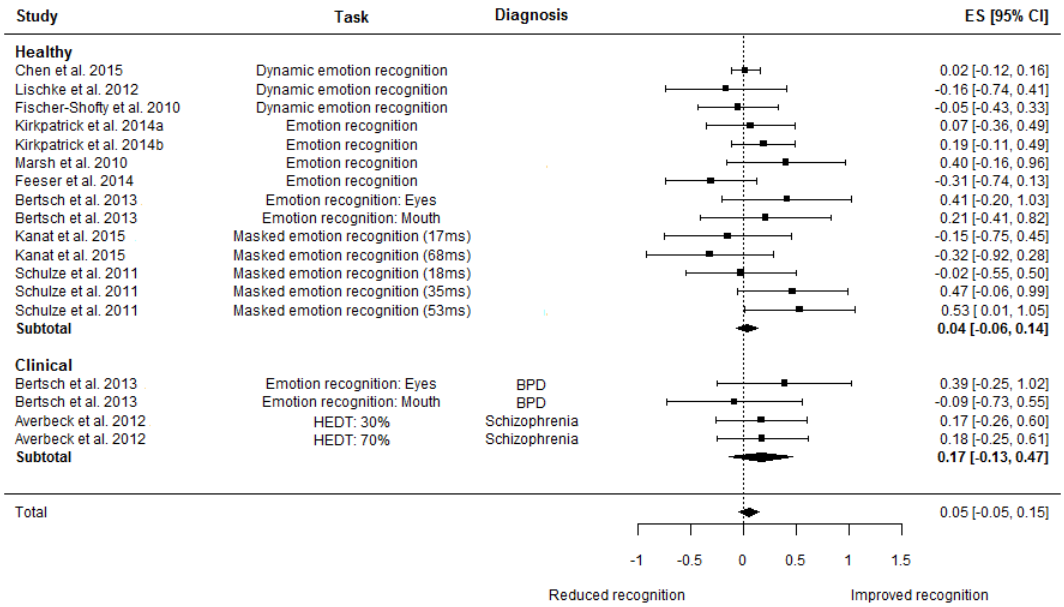
Figure 12. Effect of oxytocin vs. placebo on emotion expression. Positive effect sizes indicate increased emotion expression following oxytocin administration; negative effect sizes indicate reduced emotion expression following oxytocin administration. IAPS = International affective picture system; AN = anorexia nervosa; BPD = Borderline personality disorder.

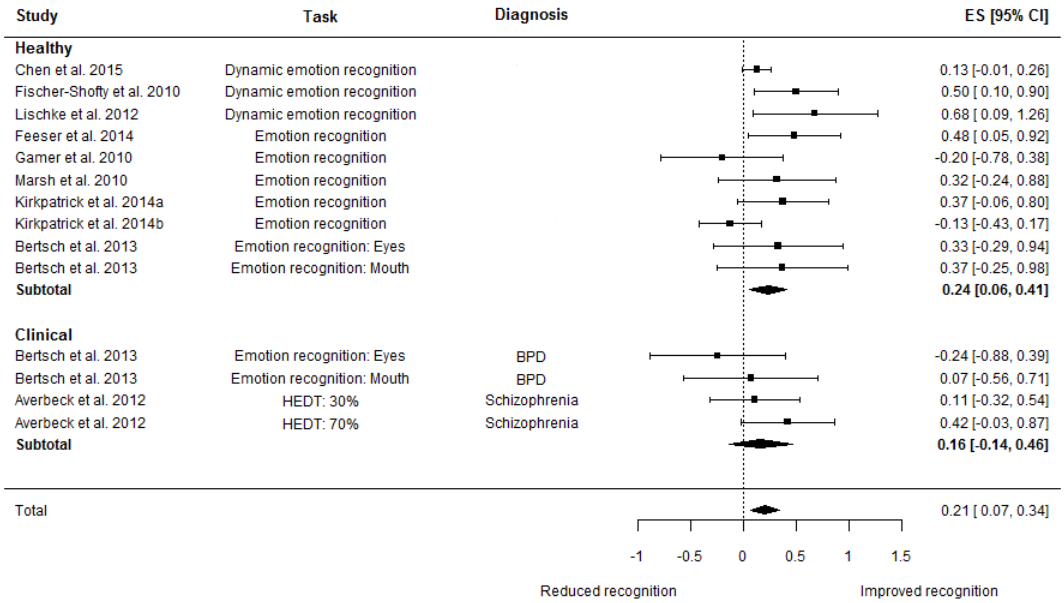

Identification
Records identified through database searching (N = 2408)
Screening
Records after duplicates removed (N = 1358)
Records screened (N = 1358)
Records excluded (N = 1294)
Eligibility
Full text articles assessed for eligibility (N = 64)
Full text articles excluded for the following reasons:
• 17 did not assess recognition accuracy / recognition sensitivity / theory of mind accuracy
• 2 used repeat administration
• 12 studies did not report relevant data
Included
Studies included in meta-analyses (N = 33)
- 1 study was identified as an outlier
- Final sample N = 32

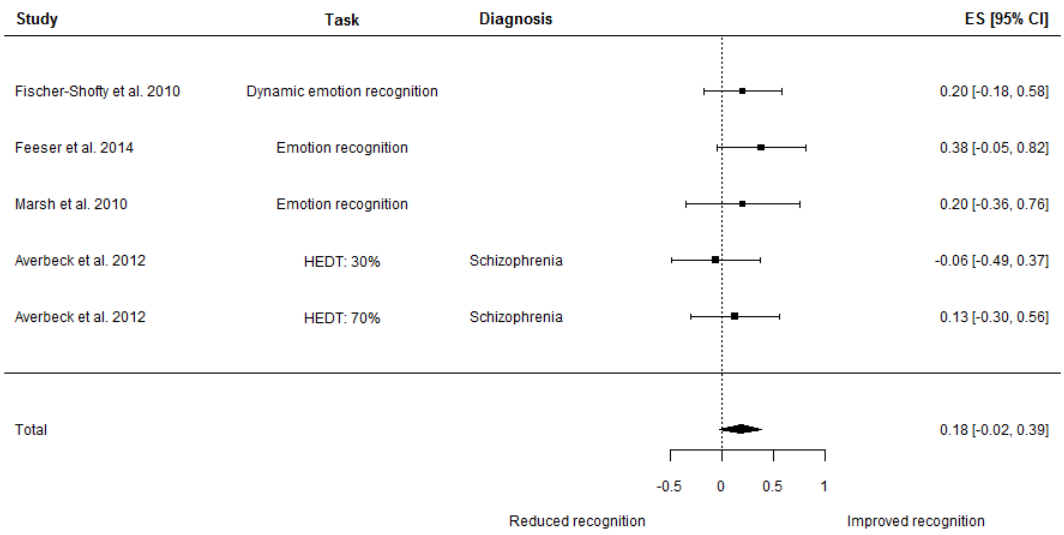

Identification
Records identified through database searching (N = 2880)
Screening
Records after duplicates removed (N = 1709)
Records screened (N = 1709)
Records excluded (N = 1699)
Eligibility
Full text articles assessed for eligibility (N = 10)
Full text articles excluded for the following reasons:
• 4 did not assess positive and negative facial expressivity
• 2 included children
Included
Studies included in meta-analysis (N = 4)

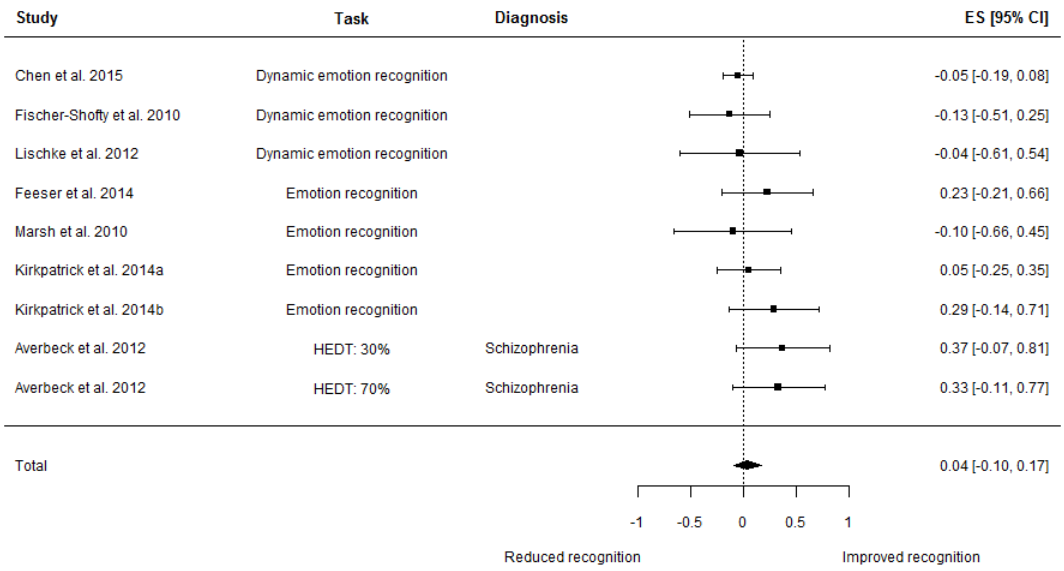

Study
Task
Diagnosis
ES [95% CI]
Healthy
Domes et al. 2007 RMET 0.47 [0.01, 0.93]
Feeser et al. 2015 RMET 0.62 [0.14, 1.09]
Luminet et al. 2011 RMET 0.30 [-0.21, 0.81]
Radke & de Bruijn 2015 RMET 0.09 [-0.31, 0.49]
Riem et al. 2014 RMET -0.02 [-0.57, 0.54]
Woolley et al. 2014 RMET 0.14 [-0.21, 0.50]
Woolley et al. 2016 RMET -0.06 [-0.40, 0.28]
Leppanen et al. 2017 RMET -0.21 [-0.58, 0.16]
Woolley et al. 2014 TASIT: Social -0.38 [-0.73, -0.02]
Woolley et al. 2014 TASIT: Emotion 0.04 [-0.33, 0.40]
Cardoso et al. 2014 MSCEIT: Faces -0.61 [-1.05, -0.18]
Cardoso et al. 2014 MSCEIT: Changes 0.08 [-0.35, 0.52]
Cardoso et al. 2014 MSCEIT: Blends 0.05 [-0.39, 0.48]
Subtotal 0.10 [-0.17, 0.37]
Clinical
Leppanen et al. 2017 RMET AN 0.10 [-0.26, 0.46]
Mitchell et al. 2016 RMET Alcohol dependence 0.24 [-0.12, 0.60]
Aoki et al. 2014 Sally Anne task ASD 0.54 [0.07, 1.01]
MacDonald et al. 2013 RMET Depression 0.50 [-0.00, 1.00]
Jesso et al. 2011 RMET FTD -0.46 [-0.92, 0.01]
Woolley et al. 2016 RMET Opioid dependence -0.51 [-0.88, -0.15]
Guastella et al. 2015 RMET Schizophrenia 0.39 [-0.05, 0.84]
Woolley et al. 2014 RMET Schizophrenia -0.02 [-0.38, 0.35]
Woolley et al. 2014 TASIT: Social Schizophrenia 0.35 [-0.01, 0.72]
Woolley et al. 2014 TASIT: Emotion Schizophrenia -0.02 [-0.39, 0.36]
Subtotal -0.07 [-0.72, 0.57]
Total 0.09 [-0.06, 0.24]
-1.5
-1
-0.5
0
0.5
1
1.5
Reduced theory of mind
Improved theory of mind

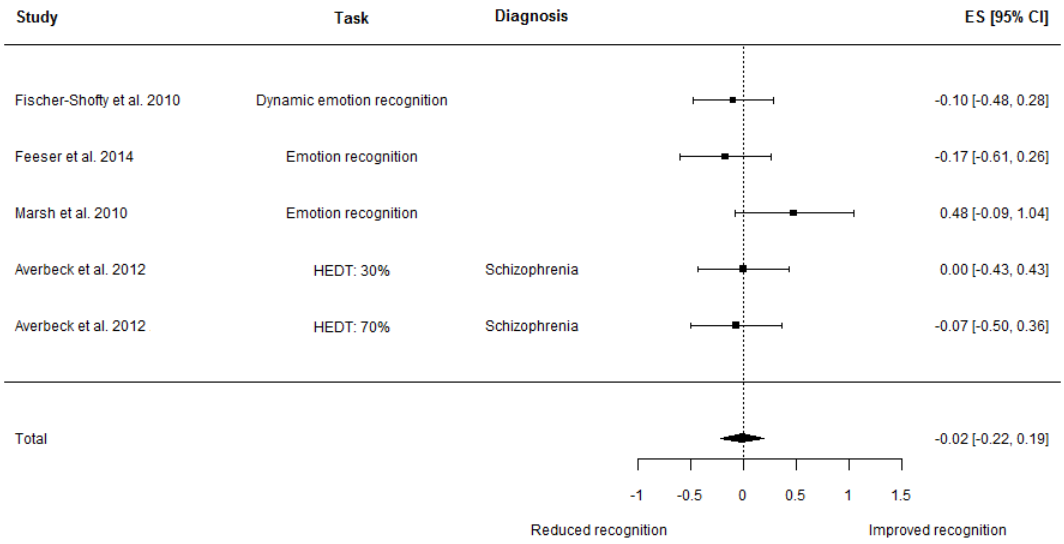

Study
Task
Diagnosis
ES [95% CI]
Healthy
Chen et al. 2015 Dynamic emotion recognition 0.01 [-0.13, 0.15]
Fischer-Shofty et al. 2010 Dynamic emotion recognition 0.15 [-0.23, 0.53]
Lischke et al. 2012 Dynamic emotion recognition 0.10 [-0.47, 0.67]
Kirkpatrick et al. 2014a Dynamic emotion recognition -0.05 [-0.48, 0.37]
Kirkpatrick et al. 2014b Dynamic emotion recognition 0.27 [-0.03, 0.57]
Kirsch et al. 2005 Emotion matching 0.07 [-0.43, 0.58]
Koch et al. 2016 Emotion matching -0.09 [-0.40, 0.22]
Marsh et al. 2010 Emotion recognition 0.36 [-0.19, 0.92]
Feeser et al. 2014 Emotion recognition 0.60 [0.16, 1.04]
Gamer et al. 2010 Emotion recognition 0.06 [-0.52, 0.63]
Campbell et al. 2014a Emotion recognition 0.07 [-0.60, 0.74]
Campbell et al. 2014b Emotion recognition -0.01 [-0.69, 0.66]
Campbell et al. 2014c Emotion recognition -0.16 [-0.83, 0.51]
Campbell et al. 2014d Emotion recognition 0.77 [0.07, 1.46]
Domes et al. 2014 Emotion recognition: Eyes 0.31 [-0.23, 0.84]
Bertsch et al. 2013 Emotion recognition: Eyes 0.27 [-0.34, 0.89]
Domes et al. 2014 Emotion recognition: Mouth -0.26 [-0.79, 0.27]
Bertsch et al. 2013 Emotion recognition: Mouth 0.23 [-0.39, 0.84]
Kanat et al. 2015 Masked emotion recognition (17ms) -0.05 [-0.65, 0.55]
Kanat et al. 2015 Masked emotion recognition (68ms) -0.09 [-0.69, 0.50]
Schulze et al. 2011 Masked emotion recognition (18ms) 0.29 [-0.24, 0.82]
Schulze et al. 2011 Masked emotion recognition (35ms) 0.64 [0.11, 1.17]
Schulze et al. 2011 Masked emotion recognition (53ms) 0.60 [0.08, 1.13]
Subtotal 0.13 [0.02, 0.24]
Clinical
Domes et al. 2014 Emotion recognition: Eyes ASD 0.58 [0.02, 1.15]
Domes et al. 2014 Emotion recognition: Mouth ASD 0.53 [-0.03, 1.09]
Bertsch et al. 2013 Emotion recognition: Eyes BPD -0.02 [-0.66, 0.61]
Bertsch et al. 2013 Emotion recognition: Mouth BPD -0.14 [-0.77, 0.50]
Koch et al. 2016 Emotion matching PTSD 0.36 [0.03, 0.70]
Guastella et al. 2015 Emotion recognition Schizophrenia -0.40 [-0.84, 0.05]
Averbeck et al. 2012 HEDT: 30% Schizophrenia 0.58 [0.12, 1.05]
Averbeck et al. 2012 HEDT: 70% Schizophrenia 0.58 [0.12, 1.05]
Guastella et al. 2015 DANVA: faces Schizophrenia 0.13 [-0.30, 0.56]
Subtotal 0.27 [-0.02, 0.57]
Total 0.18 [0.06, 0.29]
-1
-0.5
0
0.5
1
1.5
Reduced recognition
Improved recognition

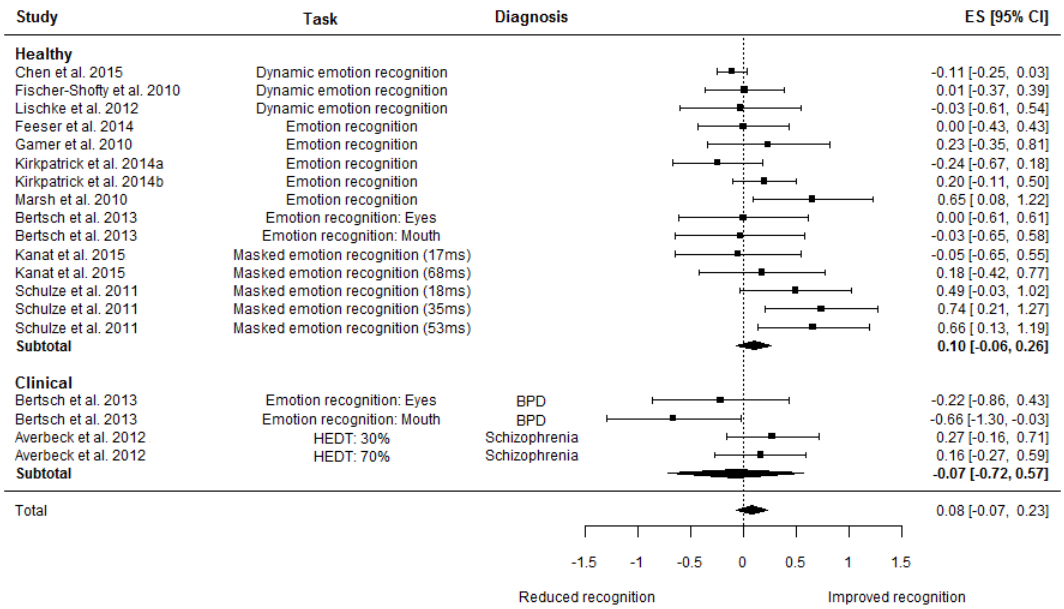

Study
Task
Diagnosis
ES [95% CI]
Healthy
Chen et al. 2015 Dynamic emotion recognition 0.02 [-0.12, 0.16]
Lischke et al. 2012 Dynamic emotion recognition -0.16 [-0.74, 0.41]
Fischer-Shofty et al. 2010 Dynamic emotion recognition -0.05 [-0.43, 0.33]
Kirkpatrick et al. 2014a Emotion recognition 0.07 [-0.36, 0.49]
Kirkpatrick et al. 2014b Emotion recognition 0.19 [-0.11, 0.49]
Marsh et al. 2010 Emotion recognition 0.40 [-0.16, 0.96]
Feeser et al. 2014 Emotion recognition -0.31 [-0.74, 0.13]
Bertsch et al. 2013 Emotion recognition: Eyes 0.41 [-0.20, 1.03]
Bertsch et al. 2013 Emotion recognition: Mouth 0.21 [-0.41, 0.82]
Kanat et al. 2015 Masked emotion recognition (17ms) -0.15 [-0.75, 0.45]
Kanat et al. 2015 Masked emotion recognition (68ms) -0.32 [-0.92, 0.28]
Schulze et al. 2011 Masked emotion recognition (18ms) -0.02 [-0.55, 0.50]
Schulze et al. 2011 Masked emotion recognition (35ms) 0.47 [-0.06, 0.99]
Schulze et al. 2011 Masked emotion recognition (53ms) 0.53 [0.01, 1.05]
Subtotal 0.04 [-0.06, 0.14]
Clinical
Bertsch et al. 2013 Emotion recognition: Eyes BPD 0.39 [-0.25, 1.02]
Bertsch et al. 2013 Emotion recognition: Mouth BPD -0.09 [-0.73, 0.55]
Averbeck et al. 2012 HEDT: 30% Schizophrenia 0.17 [-0.26, 0.60]
Averbeck et al. 2012 HEDT: 70% Schizophrenia 0.18 [-0.25, 0.61]
Subtotal 0.17 [-0.13, 0.47]
Total 0.05 [-0.05, 0.15]
-1
-0.5
0
0.5
1
1.5
Reduced recognition
Improved recognition

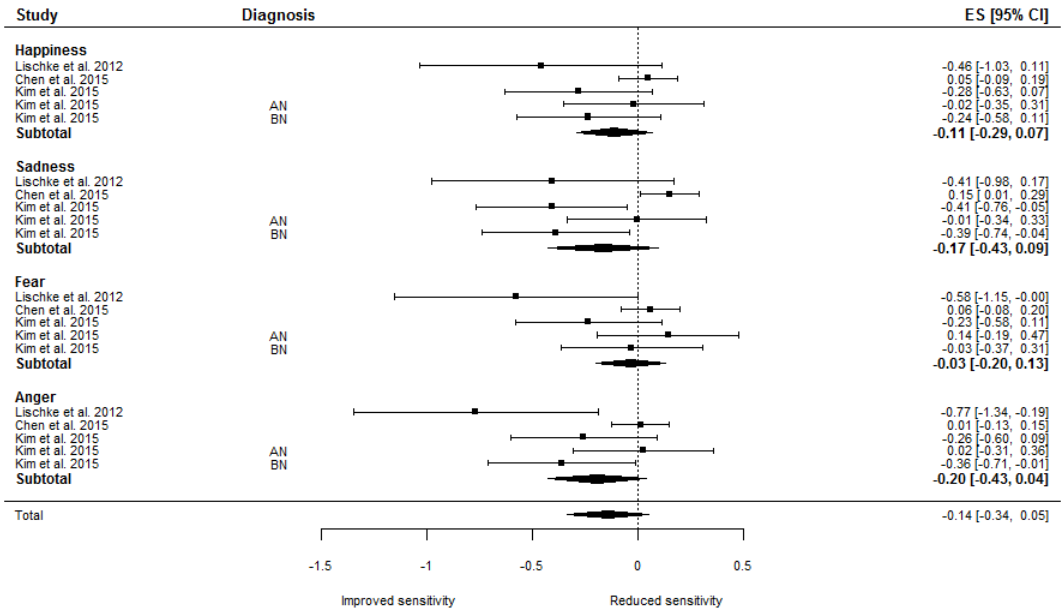

Study
Task
Diagnosis
ES [95% CI]
Healthy
Chen et al. 2015 Dynamic emotion recognition 0.13 [-0.01, 0.26]
Fischer-Shofty et al. 2010 Dynamic emotion recognition 0.50 [0.10, 0.90]
Lischke et al. 2012 Dynamic emotion recognition 0.68 [0.09, 1.26]
Feeser et al. 2014 Emotion recognition 0.48 [0.05, 0.92]
Gamer et al. 2010 Emotion recognition -0.20 [-0.78, 0.38]
Marsh et al. 2010 Emotion recognition 0.32 [-0.24, 0.88]
Kirkpatrick et al. 2014a Emotion recognition 0.37 [-0.06, 0.80]
Kirkpatrick et al. 2014b Emotion recognition -0.13 [-0.43, 0.17]
Bertsch et al. 2013 Emotion recognition: Eyes 0.33 [-0.29, 0.94]
Bertsch et al. 2013 Emotion recognition: Mouth 0.37 [-0.25, 0.98]
Subtotal 0.24 [0.06, 0.41]
Clinical
Bertsch et al. 2013 Emotion recognition: Eyes BPD -0.24 [-0.88, 0.39]
Bertsch et al. 2013 Emotion recognition: Mouth BPD 0.07 [-0.56, 0.71]
Averbeck et al. 2012 HEDT: 30% Schizophrenia 0.11 [-0.32, 0.54]
Averbeck et al. 2012 HEDT: 70% Schizophrenia 0.42 [-0.03, 0.87]
Subtotal 0.16 [-0.14, 0.46]
Total 0.21 [0.07, 0.34]
-1
-0.5
0
0.5
1
1.5
Reduced recognition
Improved recognition


Study
Task
Diagnosis
ES [95% CI]
Fischer-Shofty et al. 2010
Dynamic emotion recognition
0.20 [-0.18, 0.58]
Feeser et al. 2014
Emotion recognition
0.38 [-0.05, 0.82]
Marsh et al. 2010
Emotion recognition
0.20 [-0.36, 0.76]
Averbeck et al. 2012
HEDT: 30%
Schizophrenia
-0.06 [-0.49, 0.37]
Averbeck et al. 2012
HEDT: 70%
Schizophrenia
0.13 [-0.30, 0.56]
Total
0.18 [-0.02, 0.39]
-0.5
0
0.5
1
Reduced recognition
Improved recognition


Study
Task
Diagnosis
ES [95% CI]
Chen et al. 2015
Dynamic emotion recognition
-0.05 [-0.19, 0.08]
Fischer-Shofty et al. 2010
Dynamic emotion recognition
-0.13 [-0.51, 0.25]
Lischke et al. 2012
Dynamic emotion recognition
-0.04 [-0.61, 0.54]
Feeser et al. 2014
Emotion recognition
0.23 [-0.21, 0.66]
Marsh et al. 2010
Emotion recognition
-0.10 [-0.66, 0.45]
Kirkpatrick et al. 2014a
Emotion recognition
0.05 [-0.25, 0.35]
Kirkpatrick et al. 2014b
Emotion recognition
0.29 [-0.14, 0.71]
Averbeck et al. 2012
HEDT: 30%
Schizophrenia
0.37 [-0.07, 0.81]
Averbeck et al. 2012
HEDT: 70%
Schizophrenia
0.33 [-0.11, 0.77]
Total
0.04 [-0.10, 0.17]
-1
-0.5
0
0.5
1
Reduced recognition
Improved recognition


Study
Task
Diagnosis
ES [95% CI]
Fischer-Shofty et al. 2010
Dynamic emotion recognition
-0.10 [-0.48, 0.28]
Feeser et al. 2014
Emotion recognition
-0.17 [-0.61, 0.26]
Marsh et al. 2010
Emotion recognition
0.48 [-0.09, 1.04]
Averbeck et al. 2012
HEDT: 30%
Schizophrenia
0.00 [-0.43, 0.43]
Averbeck et al. 2012
HEDT: 70%
Schizophrenia
-0.07 [-0.50, 0.36]
Total
-0.02 [-0.22, 0.19]
-1
-0.5
0
0.5
1
1.5
Reduced recognition
Improved recognition


Study
Task
Diagnosis
ES [95% CI]
Healthy
Chen et al. 2015 Dynamic emotion recognition -0.11 [-0.25, 0.03]
Fischer-Shofty et al. 2010 Dynamic emotion recognition 0.01 [-0.37, 0.39]
Lischke et al. 2012 Dynamic emotion recognition -0.03 [-0.61, 0.54]
Feeser et al. 2014 Emotion recognition 0.00 [-0.43, 0.43]
Gamer et al. 2010 Emotion recognition 0.23 [-0.35, 0.81]
Kirkpatrick et al. 2014a Emotion recognition -0.24 [-0.67, 0.18]
Kirkpatrick et al. 2014b Emotion recognition 0.20 [-0.11, 0.50]
Marsh et al. 2010 Emotion recognition 0.65 [0.08, 1.22]
Bertsch et al. 2013 Emotion recognition: Eyes 0.00 [-0.61, 0.61]
Bertsch et al. 2013 Emotion recognition: Mouth -0.03 [-0.65, 0.58]
Kanat et al. 2015 Masked emotion recognition (17ms) -0.05 [-0.65, 0.55]
Kanat et al. 2015 Masked emotion recognition (68ms) 0.18 [-0.42, 0.77]
Schulze et al. 2011 Masked emotion recognition (18ms) 0.49 [-0.03, 1.02]
Schulze et al. 2011 Masked emotion recognition (35ms) 0.74 [0.21, 1.27]
Schulze et al. 2011 Masked emotion recognition (53ms) 0.66 [0.13, 1.19]
Subtotal 0.10 [-0.06, 0.26]
Clinical
Bertsch et al. 2013 Emotion recognition: Eyes BPD -0.22 [-0.86, 0.43]
Bertsch et al. 2013 Emotion recognition: Mouth BPD -0.66 [-1.30, -0.03]
Averbeck et al. 2012 HEDT: 30% Schizophrenia 0.27 [-0.16, 0.71]
Averbeck et al. 2012 HEDT: 70% Schizophrenia 0.16 [-0.27, 0.59]
Subtotal -0.07 [-0.72, 0.57]
Total 0.08 [-0.07, 0.23]
-1.5
-1
-0.5
0
0.5
1
1.5
Reduced recognition
Improved recognition


Study
Diagnosis
ES [95% CI]
Happiness
Lischke et al. 2012 -0.46 [-1.03, 0.11]
Chen et al. 2015 0.05 [-0.09, 0.19]
Kim et al. 2015 -0.28 [-0.63, 0.07]
Kim et al. 2015 AN -0.02 [-0.35, 0.31]
Kim et al. 2015 BN -0.24 [-0.58, 0.11]
Subtotal -0.11 [-0.29, 0.07]
Sadness
Lischke et al. 2012 -0.41 [-0.98, 0.17]
Chen et al. 2015 0.15 [0.01, 0.29]
Kim et al. 2015 -0.41 [-0.76, -0.05]
Kim et al. 2015 AN -0.01 [-0.34, 0.33]
Kim et al. 2015 BN -0.39 [-0.74, -0.04]
Subtotal -0.17 [-0.43, 0.09]
Fear
Lischke et al. 2012 -0.58 [-1.15, -0.00]
Chen et al. 2015 0.06 [-0.08, 0.20]
Kim et al. 2015 -0.23 [-0.58, 0.11]
Kim et al. 2015 AN 0.14 [-0.19, 0.47]
Kim et al. 2015 BN -0.03 [-0.37, 0.31]
Subtotal -0.03 [-0.20, 0.13]
Anger
Lischke et al. 2012 -0.77 [-1.34, -0.19]
Chen et al. 2015 0.01 [-0.13, 0.15]
Kim et al. 2015 -0.26 [-0.60, 0.09]
Kim et al. 2015 AN 0.02 [-0.31, 0.36]
Kim et al. 2015 BN -0.36 [-0.71, -0.01]
Subtotal -0.20 [-0.43, 0.04]
Total -0.14 [-0.34, 0.05]
-1.5
-1
-0.5
0
0.5
Improved sensitivity
Reduced sensitivity

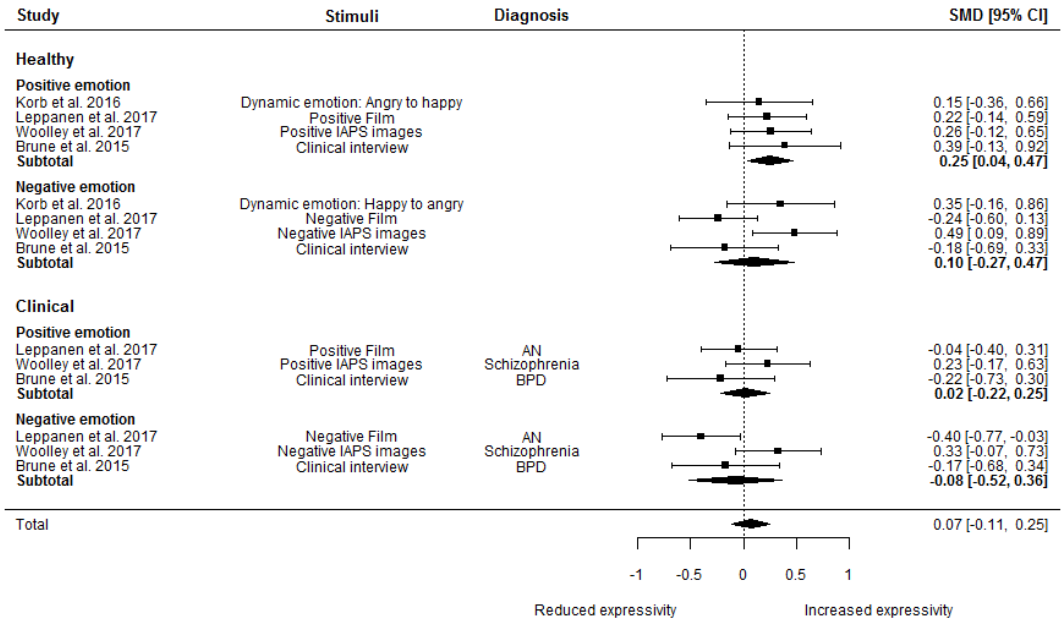

Study
Stimuli
Diagnosis
SMD [95% CI]
Healthy
Positive emotion
Korb et al. 2016 Dynamic emotion: Angry to happy 0.15 [-0.36, 0.66]
Leppanen et al. 2017 Positive Film 0.22 [-0.14, 0.59]
Woolley et al. 2017 Positive IAPS images 0.26 [-0.12, 0.65]
Brune et al. 2015 Clinical interview 0.39 [-0.13, 0.92]
Subtotal 0.25 [0.04, 0.47]
Negative emotion
Korb et al. 2016 Dynamic emotion: Happy to angry 0.35 [-0.16, 0.86]
Leppanen et al. 2017 Negative Film -0.24 [-0.60, 0.13]
Woolley et al. 2017 Negative IAPS images 0.49 [0.09, 0.89]
Brune et al. 2015 Clinical interview -0.18 [-0.69, 0.33]
Subtotal 0.10 [-0.27, 0.47]
Clinical
Positive emotion
Leppanen et al. 2017 Positive Film AN -0.04 [-0.40, 0.31]
Woolley et al. 2017 Positive IAPS images Schizophrenia 0.23 [-0.17, 0.63]
Brune et al. 2015 Clinical interview BPD -0.22 [-0.73, 0.30]
Subtotal 0.02 [-0.22, 0.25]
Negative emotion
Leppanen et al. 2017 Negative Film AN -0.40 [-0.77, -0.03]
Woolley et al. 2017 Negative IAPS images Schizophrenia 0.33 [-0.07, 0.73]
Brune et al. 2015 Clinical interview BPD -0.17 [-0.68, 0.34]
Subtotal -0.08 [-0.52, 0.36]
Total 0.07 [-0.11, 0.25]
-1
-0.5
0
0.5
1
Reduced expressivity
Increased expressivity

Table 1. Study characteristics

Study	Dose	Design	Dose-to-task interval	Group N	Age	Task	Scale	Emotion	ES [95% CI]	Power ≥80%
Healthy										
Bertsch, et al. (2013)	26 IU	Between subjects	40 min	Healthy Oxytocin = 21 Placebo = 20	Oxytocin: 24.6 (3.9) Placebo: 24.4 (4.4)	Emotion recognition: Eyes	Accuracy	Anger	0.21 [-0.41, 0.82]	No
								Fear	0.37 [-0.25, 0.99]	
								Happiness	-0.03 [-0.65, 0.58]	
								Total	0.27 [-0.32, 0.87]	
						Emotion recognition: Mouth		Anger	0.41 [-0.21, 1.03]	
								Fear	0.33 [-0.29, 0.95]	
								Happiness	<0.01 [-0.61, 0.61]	
								Total	0.23 [-0.31, 0.77]	
Brune, et al. (2015)	24 IU	Within subjects	NR	Healthy N = 15	25.70 (6.40)	Ethological Coding System for Interviews	Expression	Affiliation	0.39 [-0.13, 0.92]	No
								Flight	-0.18 [-0.69, 0.33]	
	20 IU	Between subjects	45 min	Older female N = 34	72.07 (6.49)	Emotion recognition	Accuracy	Total	-0.16 [-0.80, 0.48]	No

Campbell, et al. (2014)				Older male N = 34					0.77 [0.13, 1.40]	
				Younger female N = 34	19.68 (1.79)				0.07 [-0.54, 0.67]	
				Younger male N = 34					-0.01 [-0.62, 0.59]	
Cardoso, et al. (2014)	24 IU	Between subjects	120 min	Healthy Oxytocin = 42 Placebo = 40	18-30	Mayer- Salovey-Caruso Emotional Intelligence Test (MSCEIT)	Accuracy	Understanding emotions (blends)	0.05 [-0.45, 0.54]	No
								Understanding emotions (changes)	0.08 [-0.42, 0.59]	
								Perceiving emotions (total faces)	-0.61 [-1.17, -0.06]	
Chen, et al. (2015)	24 IU	Within subjects	45 min	Healthy N = 203	23.5 (2.7)	Dynamic emotion recognition	Accuracy	Anger	0.02 [-0.12, 0.16]	Yes
								Fear	0.13 [-0.01, 0.26]	
								Sadness	-0.05 [-0.19, 0.08]	
								Happiness	-0.11 [-0.25, 0.03]	
								Total	0.01 [-0.13, 0.15]	
								Anger	0.01 [-0.13, 0.15]	

							Recognition sensitivity	Fear	0.06 [-0.08, 0.20]	
								Sadness	0.15 [0.01, 0.29]	
								Happiness	0.05 [-0.09, 0.19]	
Domes, et al. (2007)	24 IU	Within subjects	45 min	Healthy N = 20	24.3 (2.2)	RMET	Accuracy	Total	0.47 [0.01, 0.93]	No
Domes, et al. (2014)	24 IU	Within subjects	45 min	Healthy N = 14	23.6 (5.4)	Emotion recognition: Eyes	Accuracy	Total	0.31 [-0.23, 0.84]	No
						Emotion recognition: Mouth			-0.25 [-0.79, 0.27]	
Feeser, et al. (2014)	24 IU	Between subjects	45 min	Healthy Oxytocin = 41 Placebo = 41	37.9 (4.7)	Facial emotion recognition task	Accuracy	Anger	-0.31 [-0.50, -0.11]	No
								Fear	0.48 [0.29, 0.68]	
								Disgust	0.38 [-0.17, 0.94]	
								Sadness	0.23 [-0.33, 0.78]	
								Surprise	-0.17 [-0.73, 0.39]	
								Happiness	<0.01 [-0.19, 0.19]	
								Total	0.60 [-0.16, 1.36]	
Feeser, et al. (2015)	24 IU	Between subjects	45 min	Healthy Oxytocin = 36	Oxytocin = 27.2 (4.9)	RMET	Accuracy	Total	0.62 [0.12, 1.12]	No

				Placebo = 35	Placebo = 28.9 (4.8)					
Fischer-Shofty, et al. (2010)	24 IU	Within subjects	45 min	Healthy N = 27	26.93 (3.51)	Facial emotion recognition task	Accuracy	Anger	-0.05 [-0.43, 0.33]	No
								Fear	0.50 [0.10, 0.90]	
								Disgust	0.20 [-0.18, 0.58]	
								Sadness	-0.13 [-0.51, 0.25]	
								Surprise	-0.10 [-0.48, 0.28]	
								Happiness	0.01 [-0.37, 0.39]	
								Total	0.15 [-0.23, 0.53]	
Gamer, et al. (2010)	24 IU	Between subjects	45 min	Healthy Oxytocin = 23 Placebo = 23	25.0 (3.7)	Facial emotion recognition task	Accuracy	Fear	-0.20 [-0.78, 0.38]	No
								Happiness	0.23 [-0.35, 0.81]	
								Total	0.06 [-0.52, 0.63]	
Kanat, et al. (2015)	24IU	Between subjects	45 min	Healthy Oxytocin = 21 Placebo = 22	23.64 (2.81)	Masked emotion recognition task (17ms)	Accuracy	Anger	-0.15 [-0.16, 0.96]	No
								Happiness	-0.05 [-0.65,0.55]	
								Total	-0.05 [-0.65, 0.55]	
								Anger	-0.32 [-0.74, 0.41]	

						Masked emotion recognition task (68ms)		Happiness	0.18 [-0.42, 0.77]	
								Total	-0.09 [-0.69, 0.50]	
Kim, et al. (2015)	40 IU	Within subjects	45 min	Healthy N = 33	22.64 (2.28)	Dynamic emotion recognition	Recognition sensitivity	Anger	-0.26 [-0.60, 0.09]	No
								Fear	-0.23 [-0.58, 0.11]	
								Sadness	-0.41, -0.76, -0.05]	
								Happiness	0.03 [-0.63, 0.07]	
Kirkpatrick, et al. (2014)	20 IU	Within subjects	30 min	Healthy N = 43	24.10 (4.10)	Dynamic emotion recognition	Accuracy	Anger	0.07 [-0.23, 0.36]	Yes
								Fear	-0.16 [-0.43, 0.17]	
								Sadness	0.02 [-0.25, 0.35]	
								Happiness	-0.24 [-0.55, 0.06]	
								Total	-0.05 [-0.35, 0.25]	
	40 IU			Healthy N = 22	23.10 (3.50)			Anger	0.19 [-0.23, 0.61]	No
								Fear	0.37 [-0.06, 0.80]	
								Sadness	0.29 [-0.14, 0.71]	
								Happiness	0.20 [-0.22, 0.62]	

								Total	0.27 [-0.15, 0.70]	
Kirsch, et al. (2005)	27 IU	Within subjects	50 min	Healthy N = 15	26.70 (3.00)	Emotion matching	Accuracy	Total	0.07 [-0.43, 0.58]	No
Koch, et al. (2016)	40 IU	Within subjects	44.68 min	Trauma exposed healthy N = 40	40.00 (10.05)	Emotion matching	Accuracy	Total	-0.09 [-0.40, 0.22]	Yes
Korb, et al. (2016)	24 IU	Between subjects	56 min	Healthy Oxytocin = 30 Placebo = 30	Oxytocin = 26.10 (5.10) Placebo = 23.60 (4.10)	Dynamic emotion: Happy to angry	Expression	EMG: CS	0.35 [-0.37, 1.07]	No
						Dynamic emotion: Angry to happy	Expression	EMG: ZM	0.15 [-0.39, 0.68]	
Leppanen, et al. (2017)	40 IU	Within subjects	15 min	Healthy N = 29	26.83 (8.54)	RMET	Accuracy	Total	-0.21 [-0.58, 0.16]	No
						Evoked facial expressions to film stimuli	Expression	Happiness	0.22 [-0.14, 0.59]	
								Sadness	-0.24 [-0.61, 0.13]	
Lischke, et al. (2012)	24 IU	Between subjects	45 min	Healthy Oxytocin = 23 Placebo = 24	Oxytocin = 25.78 (3.37) Placebo = 26.38 (3.49)	Dynamic emotion recognition	Accuracy	Anger	-0.16 [-0.70, 0.37]	No
								Fear	0.68 [0.24, 1.12]	
								Sadness	-0.04 [-0.47, 0.40]	
								Happiness	-0.03 [-0.57, 0.50]	
								Total	0.10 [-0.33, 0.52]	

							Recognition sensitivity	Anger	-0.77 [-1.25, -0.29]	
								Fear	-0.58 [-1.06, -0.09]	
								Sadness	-0.41 [-0.60, -0.21]	
								Happiness	-0.46 [-1.05, 0.13]	
								Total	-0.55 [-1.37, 0.03]*	
Luminet, et al. (2011)	32 IU	Between subjects	45 min	Healthy Oxytocin = 30 Placebo = 30	21.08 (2.13)	RMET	Accuracy	Total	0.30 [-0.32, 0.93]	No
Marsh, et al. (2010)	24 IU	Between subjects	35 min	Healthy Oxytocin = 24 Placebo = 26	Oxytocin = 26.20 (4.90) Placebo = 26.60 (5.00)	Facial emotion recognition task	Accuracy	Anger	0.40 [-0.12, 0.93]	No
								Fear	0.32 [-0.28, 0.91]	
								Disgust	0.20 [-0.40, 0.80]	
								Sadness	-0.10 [-0.69, 0.49]	
								Surprise	0.48 [-0.07, 1.02]	
								Happiness	0.65 [0.12, 1.18]	
								Total	0.36 [-0.16, 0.89]	
Radke and de	24 IU	Within subjects	50-65 min	Healthy N = 24	21.50 (1.90)	RMET	Accuracy	Total	0.09 [-0.31, 0.49]	No

Bruijn (2015)										
Riem, et al. (2014)	16 IU	Between subjects	60 min	Healthy Oxytocin = 25 Placebo = 25	19.62 (1.47)	RMET	Accuracy	Total	-0.02 [-0.58, 0.55]	No
Schulze, et al. (2011)	24 IU	Between subjects	45 min	Healthy Oxytocin = 28 Placebo = 28	24.18 (3.12)	Masked emotion recognition (mask: 18ms)	Accuracy	Anger	0.53 [-0.03, 1.09]	No
								Happiness	0.66 [0.09, 1.23]	
								Total	0.29 [-0.33, 0.90]	
						Masked emotion recognition (mask: 35ms)		Anger	0.47 [-0.11, 1.04]	
								Happiness	0.74 [0.16, 1.31]	
								Total	0.64 [-0.11, 1.38]	
						Masked emotion recognition (mask: 53ms)		Anger	-0.02 [-0.45, 0.40]	
								Happiness	0.49 [0.07, 0.92]	
								Total	0.60 [-0.14, 1.35]	
Woolley, et al. (2016)	40 IU	Within subjects	45 min	Healthy N = 33	51.91 (7.35)	RMET	Accuracy	Total	-0.06 [-0.40, 0.28]	No
	40 IU	Within subjects	30 min	Healthy N = 31	42.50 (14.10)	TASIT: Emotion Evaluation Test	Accuracy	Total	0.04 [-0.32, 0.39]	No

Woolley, et al. (2014)						TASIT: Social Inference Enriched (feel)			-0.38 [-0.74, 0.01]	
						RMET			0.14 [-0.21, 0.50]	
Woolley, et al. (2017)	40 IU	Within subjects	NR	Healthy N = 27	42.00 (13.70)	Evoked facial expressions to IAPS photos	Expression	Positive	0.26 [-0.12, 0.65]	No
								Negative	0.49 [0.09, 0.89]	
Xu, et al. (2015)	40 IU	Between subjects	45 min	Healthy Oxytocin = 29 Placebo = 31	Oxytocin = 23.40 (0.30) Placebo = 22.90 (0.30)	RSVP task	Accuracy	Negative	0.76 [0.23, 1.28]*	No
								Happiness	3.14 [2.39, 3.90][1]	
								Neutral	4.56 [3.60, 5.52]*	
								Total	2.96 [2.23, 3.69]*	
Clinical populations										
Aoki, et al. (2015)	24 IU	Within subjects	40 min	ASD N = 20	30.8 (6.00)	Sally Anne: Social-emotional scale	Accuracy	Total	0.54 [0.07, 1.01]	No
Averbeck (2012)	24 IU	Within subjects	50 min	Schizophrenia N = 21	38.2 (1.8)	Hexagon emotion discrimination task: morphed faces 30% intensity	Accuracy	Anger	0.17 [-0.26, 0.60]	No
								Fear	0.11 [-0.32, 0.54]	
								Disgust	-0.06 [-0.49, 0.37]	
								Sadness	0.37 [-0.07, 0.81]	
								Surprise	<0.01 [-0.43, 0.43]	

								Happiness	0.27 [-0.16, 0.71]	
								Total	0.58 [0.12, 1.05]	
						Hexagon emotion discrimination task: morphed faces 70% intensity		Anger	0.18 [-0.25, 0.61]	
								Fear	0.42 [-0.03, 0.87]	
								Disgust	0.13 [-0.30, 0.56]	
								Sadness	0.33 [-0.11, 0.77]	
								Surprise	-0.07 [-0.50, 0.36]	
								Happiness	0.16 [-0.27, 0.59]	
								Total	0.58 [0.12, 1.05]	
Bertsch, et al. (2013)	26 IU	Between subjects	40 min	BPD Oxytocin = 19 Placebo = 19	Oxytocin: 23.2 (5.3) Placebo: 24.9 (5.5)	Emotion recognition: Eyes	Accuracy	Anger	-0.09 [-0.73, 0.55]	No
								Fear	0.07 [-0.56, 0.71]	
								Happiness	-0.66 [-1.32, -0.01]	
								Total	-0.02 [-0.49, 0.44]	
						Emotion recognition: Mouth		Anger	0.39 [-0.25, 1.03]	
								Fear	-0.24 [-0.88, 0.39]	

								Happiness	-0.22 [-0.85, 0.42]	
								Total	-0.14 [-0.85, 0.58]	
Brune, et al. (2015)	24 IU	Within subjects	NR	BPD N = 15	27.50 (7.30)	Ethological Coding System for Interviews	Expression	Affiliation	-0.22 [-0.73, 0.30]	No
								Flight	-0.17 [-0.68, 0.34]	
Domes, et al. (2014)	24 IU	Within subjects	45 min	ASD N = 14	24.0 (6.9)	Emotion recognition: Eyes	Accuracy	Total	0.58 [0.02, 1.15]	No
						Emotion recognition: Mouth			0.53 [-0.03, 1.09]	
Guastella, et al. (2015)	24 IU	Within subjects	45 min	Schizophrenia N = 21	37.42 (11.14)	Diagnostic Analysis of Non-Verbal Accuracy: Faces	Accuracy	Total	0.13 [-0.30, 0.56]	No
						Facial Expressions of Emotions Task			-0.40 [-0.84, 0.05]	
						RMET			0.39 [-0.05, 0.84]	
Jesso, et al. (2011)	24 IU	Within subjects	20 min	FTD N = 20	64.4 (7.40)	RMET	Accuracy	Total	-0.46 [-0.92, 0.01]	No
Kim, et al. (2015)	40 IU	Within subjects	45 min	AN N = 35	21.97 (8.41)	Dynamic emotion recognition	Recognition sensitivity	Anger	0.02 [-0.31, 0.36]	Yes
								Fear	0.14 [-0.19, 0.47]	
								Sadness	-0.01 [-0.34, 0.33]	
								Happiness	0.03 [-0.35, 0.31]	

				BN N = 34	23.03 (5.17)			Anger	-0.36 [-0.71, -0.01]	Yes
								Fear	-0.03 [-0.37, 0.31]	
								Sadness	-0.39 [-0.74, -0.04]	
								Happiness	0.03 [-0.58, 0.11]	
Koch, et al. (2016)	40 IU	Within subjects	44.68 min	PTSD N = 36	39.93 (9.58)	Emotion matching	Accuracy	Total	0.36 [0.03, 0.70]	Yes
Leppanen, et al. (2017)	40 IU	Within subjects	15 min	AN N = 30	26.2 (6.82)	RMET	Accuracy	Total	0.10 [-0.26, 0.46]	No
						Evoked facial expressions to film stimuli	Expression	Happiness	-0.04 [-0.40, 0.31]	
								Sadness	-0.40 [-0.77, -0.03]	
MacDonald, et al. (2013)	40 IU	Within subjects	75 min	Depression N = 17	43.65 (12.20)	RMET	Accuracy	Total	0.50 [-0.01, 1.00]	No
Mitchell, et al. (2016)	40 IU	Within subjects	30 min	Alcohol dependence N = 32	28.90 (7.15)	RMET	Accuracy	Total	0.24 [-0.12, 0.60]	No
Woolley, et al. (2016)	40 IU	Within subjects	45 min	Opioid dependence N = 33	57.97 (8.88)	RMET	Accuracy	Total	-0.51 [-0.88, -0.15]	No
Woolley, et al. (2014)	40 IU	Within subjects	30 min	Schizophrenia N = 29	44.60 (10.70)	TASIT: Emotion Evaluation Test	Accuracy	Total	-0.02 [-0.38, 0.35]	No
						TASIT: Social Inference Enriched (feel)			0.35 [-0.02, 0.73]	

						RMET			-0.02 [-0.38, 0.35]	
Woolley, et al. (2017)	40 IU	Within subjects	NR	Schizophrenia N = 25	43.20 (11.00)	Evoked facial expressions to IAPS photos	Expression	Positive	0.23 [-0.17, 0.63]	No
								Negative	0.33 [-0.07, 0.73]	

[1] Influential outlier. Excluded from meta-analysis.

* Not included in meta-analysis.

ES = effect size; IU = international unit; NR = not reported; AN = anorexia nervosa, ASD = autism spectrum disorder, BN = bulimia nervosa, BPD = borderline personality disorder, PTSD = posttraumatic stress disorder RMET = reading the mind in the eyes test; TASIT = The Awareness of Social Inference Test; IAPS = International Affective Picture System. “No” in the power column indicates the study did not reach the sample size requirement for adequate statistical power and “Yes” in the power column indicates the study reached the sample size requirement for adequate statistical power.