Affective cognition in eating disorders: a systematic review and meta-analysis of the performance on the "Reading the Mind in the Eyes" Test

Antonio Preti[1] · Sara Siddi[2] · Enrica Marzola[1] · Giovanni Abbate Daga[1]

Abstract

Background The Reading the Mind in the Eyes Test (RMET) is listed in the National Institute of Mental Health's Research Domain Criteria as a tool apt to measure the understanding of others' mental states. People diagnosed with anorexia nervosa (AN) showed poorer performances on the RMET than healthy controls. Less data are available concerning other eating disorders.

Methods Systematic review of four major databases from inception to July 15, 2021 following the PRISMA guidelines. Meta-analysis of cross-sectional observational studies comparing the scores of the RMET between patients with eating disorders and age- and gender matched control groups.

Results Out of 21 studies, we retrieved 29 independent samples of patients diagnosed with an eating disorder. Patients with active AN ($n = 580$) showed worse performances on the RMET than controls ($n = 1019$). Year of publication accounted for 61% of the (substantial: $I^2 = 81\%$) heterogeneity in the meta-analysis. Earlier studies were more likely to find worse performances on the RMET of patients with active AN than later studies. Patients with bulimia nervosa ($n = 185$) performed poorly as compared to controls ($n = 249$), but the results were not statistically significant on the random-effect model. Obese patients with binge-eating disorder ($n = 54$) did not differ on the RMET from obese controls ($n = 52$). Patients with eating disorder not otherwise specified ($n = 57$) showed minimal differences compared to controls ($n = 96$). Study quality was good in six studies only.

Conclusions Patients with eating disorders do not suffer from an impaired understanding of others' mental states, except for a still-to-be-identified subgroup of patients with active AN.

Level of evidence I, systematic review and meta-analysis.

Introduction

Current neuroscience posits that there are brain systems dedicated to the perception and interpretation of others' actions [1, 2]. Within this framework, the US National Institute of Mental Health's (NIMH) Research Domain Criteria (RdoC) have identified the understanding of mental states as a core domain in the perception and understanding of others. This domain, akin to the concept of "Theory of Mind" (ToM; [3]), covers all the processes that are involved in the judgment and/or attribution of mental states to other animated entities so as to predict or interpret their behaviors [4].

The Reading the Mind in the Eyes Test (hereinafter the RMET) is listed in the NIMH RdoC as a tool apt to measure the understanding of others' mental states. In its final format, the RMET consists of 36 pictures of the eyes region on the face drawn from magazines and newspapers [5]. The candidate is presented with the pictures alongside four words that are supposed to describe what the pictured person is thinking or feeling. One word is the expected correct designator and

✉ Antonio Preti
antonio.preti@unito.it; apreti@tin.it

1 Eating Disorders Unit, Department of Neuroscience "Rita Levi Montalcini", University of Turin, via Cherasco 15, 10126 Turin, Italy

2 Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, Institut de Recerca Sant Joan de Déu, Sant Boi de Llobregat, Universitat de Barcelona, Dr. Antoni Pujadas, 42, 08830 Sant Boi de Llobregat, Barcelona, Spain

the other three words are foils. The candidate can consult a glossary to get a better understanding of the meaning of the words accompanying the picture. The score is the sum of all correct answers. The RMET is thought to measure a single factor [6, 7], although some authors have found a multidimensional factor structure of the RMET [8]. Performance on the RMET depends on intelligence quotient (IQ), and in particular, verbal IQ [9], and its psychometric properties vary according to the samples and version, especially its reliability and its links with self-report measures of empathy [10]. Women outperform men on the RMET [10, 11], and there is some evidence that accuracy may vary between same-race targets and other-race targets [12]. Activation in some cortical areas and related structures were seen during the execution of the RMET, in particular, the superior temporal sulcus, inferior frontal gyrus, medial prefrontal cortex, hippocampus, and cerebellum [13, 14].

The RMET was devised to assess social cognition in autistic people [5], and in fact, the task appears best suited for testing populations at the lower end of performances on the measured ability [7]. Nevertheless, over time, it has become a popular task for the assessment of emotion recognition, understanding of complex mental states, and ToM capacity in both clinical and non-clinical populations (e.g., [15, 16]). The RMET assesses the ability to recognize complex mental states as expressed by human eyes. However, most ToM tasks measure the ability of the candidate to infer non-emotional mental states, such as intentions or beliefs, from contextual information or dynamic behavioral cues. Thus, some authors questioned the specificity of the RMET as ToM task, suggesting it is principally a measure of emotion recognition [17]. Nevertheless, general consensus distinguishes between a cognitive component of ToM, which is the ability to attribute mental states to others via perspective taking, and an affective component of ToM, which is based on the ability to infer another person's emotions via facial or body expressions [18, 19]. The RMET is a reasonably apt measure of the affective component of ToM, since emotion recognition and appraisal contribute to the attribution of mental states to others. Moreover, test performances on the RMET were consistently related to the activation of the dorsomedial prefrontal cortex and the temporoparietal junction [20], which are typically related to ToM tasks [19]. Test performances on the RMET were as well related to the within-network connectivity of the right posterior superior temporal sulcus [21], and the gray matter density in the left posterior superior temporal sulcus and its functional connection with the amygdala [22], which are implied in social perception, another component of ToM [19].

Superior performances of controls over autistic people are the most replicated finding with the RMET [23]. People diagnosed with schizophrenia also show poorer performances on the RMET than healthy controls [24], with

effect sizes comparable to those observed in autistic people [25]. Conflicting results were found in samples of patients diagnosed with borderline personality disorder, with some studies finding enhanced performances compared to controls [26], and other studies finding the reverse [27–29]. Even less consistent findings were reported in patients with bipolar disorder [30–33], in those with major depressive disorder [34, 35], and in patients with obsessive–compulsive disorder [36, 37].

The impairment in the abilities involved in the RMET was related to deficits in social functioning and poor insight in schizophrenia [38, 39]. Impaired social functioning is also a core characteristic of autism spectrum disorder [40]. Individuals with anorexia nervosa (AN), too, are known to suffer from a deficit in social functioning [41–44]. They also experience a poor insight about their symptoms, sometimes so severe to reach the level of delusion [45]. Moreover, people diagnosed with AN manifest neuropsychological features similar to those that are observed in autism spectrum disorder, such as weak central coherence, cognitive inflexibility, and problem in emotion recognition [46]. Unsurprisingly, people diagnosed with AN showed poorer performances on the RMET than healthy controls, with larger effect sizes in acute patients as compared to recovered ones [47]. Social anxiety, poor social support, and interpersonal difficulties were reported also in patients with bulimia nervosa (BN; [41, 48]), who also showed reduced performances on the RMET compared to controls [47]. People with binge-eating disorder (BED) or with eating disorder not otherwise specified (EDNOS) were less often investigated on their performances on the RMET [49, 50], although both people with BED [51], and EDNOS [52], experience interpersonal difficulties. Moreover, BED and EDNOS displaying disordered eating with binge eating may evolve into obesity [53]. There is some evidence that people with obesity exhibit some degree of impairment on instruments assessing emotion recognition and ToM [54]. Thus, the investigation of affective cognition in patients diagnosed with BED and EDNOS is justified.

Conflicting results emerged in recent studies regarding the capacity to understand mental states in people diagnosed with an eating disorder. Some studies reported that patients with AN were as good as controls in responding to the RMET [55, 56], and similar findings were reported for patients diagnosed with BN [57, 58]. It is possible that the bias towards positive findings, i.e., the observation of a difference with controls—the most frequent cause of publication bias—, led to a focus on deficient performances on the RMET of people with eating disorders in earlier studies. For this reason, the role of the year of publication should be taken into account when investigating the performances on the RMET of patients with eating disorders. Another element inextricably interlaced with the year of

publication is the role played by the criteria for diagnosis in the enrollment of samples. Following its introduction, it became evident that the samples of patients diagnosed with AN or BN according to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) may include less severe cases than those included in samples that were diagnosed according to the more restrictive criteria of the DSM-IV. Indeed, cases that would have been classified under the label of EDNOS according to the DSM-IV are now diagnosed under one major category (AN, BN, BED) according to the DSM-5 [59]. Less severe cases might be less impaired on affective cognition, thus leading to fewer differences in the RMET between cases and controls in studies done according to DSM-5 criteria than in previous studies based on past DSM criteria.

Social cognition takes on an important role in behavioral disorders and may become the target of therapeutic and rehabilitative interventions [60, 61]. Thus, a reappraisal of the performances of people with eating disorders on a task aimed at measuring a core component of social cognition—the understanding of others' mental states—is worthwhile.

So far, the performances on the RMET of patients diagnosed with an eating disorder were summarized in a meta-analysis focusing on those with AN and BN, which dates back to 2016 and mostly included studies done before the DSM-5 entered into standard use in research [47]. The topic was also briefly covered in a systematic review devoted to mentalization in patients with eating disorders [62], and in a meta-analysis investigating autistic features in patients with anorexia nervosa [46]. The demonstration that a deficit of affective cognition is specific to AN and is less evident or absent in other eating disorders would support a role for it in the etiology of AN. Nevertheless, it cannot be excluded that a deficit of affective cognition in AN is merely a reflection of malnutrition. The impact of body mass index (BMI), as a proxy for the severity of AN, on the performance on the RMET might serve the purpose of investigating the role of malnutrition in AN. Differences in the performances on the RMET between samples of recovered AN and those with active AN would further corroborate a role for malnutrition.

Aims

This systematic review and meta-analysis was set out to: (a) summarize the available evidence on the performances on the RMET of people diagnosed with an eating disorder; (b) evaluate the quality of available evidence; (c) highlight current strengths and evidence gaps. We focused on studies that included either community or clinical samples irrespectively of gender or age; with performances on the RMET as an outcome; that had eating disorders as the criterion of exposure; that reported information by comparison with controls without eating disorders.

Methods

This meta-analysis was done according to the indications of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [63, 64]. We searched PubMed/Medline, the Cochrane library, the Excerpta Medica Database (EMBASE), and PsycInfo without time limitations from inception until 15 July 2021. A combination of the following key terms was used: Reading the Mind in the Eyes Test, Reading the Mind in the Eyes, RMET, AND eating disorders, anorexia nervosa, bulimia nervosa, binge-eating disorder, eating disorder not otherwise specified, other specified feeding or eating disorder.

Studies were included when they were published in a peer-reviewed journal; abstracts and unpublished theses were excluded. Indeed, there is evidence that selection bias is higher in unpublished literature than in published literature [65, 66]. Based on the knowledge of the languages of the evaluators, articles that were written in English, Spanish, French, or Italian were assessed. However, there is some evidence that a search of English language literature is enough to produce results that are similar to those that can be retrieved, with more time and effort, from reviews based on comprehensive searches free of language restrictions [65].

Two evaluators performed the search. The evaluators had 15 and 30 years of experience, respectively, in conducting research in the behavioral field, and had 6 and 10 years of experience in conducting systematic reviews and meta-analysis. The list of the retrieved articles was inspected by the two evaluators to establish whether the articles were congruent with the search criteria. Duplicates were eliminated. Discrepancies were solved by discussion. Collected articles were then thoroughly re-examined for content and their references section was scanned to identify missed studies. The same procedure was applied to the scanning of the additional sources (systematic reviews/meta-analyses). When a group published more than one report with a probable overlap of the samples, the study with the largest sample was included (Fig. 1).

Two evaluators extracted the following data from each article: author(s) and year of publication; the geographical location of the study; condition under test; criteria for diagnosis; sample size and composition by gender and age; body mass index; duration of illness; mean scores and standard deviation (SD) of the participants on the RMET by group. When the authors reported the median and the interquartile range, the median was considered equivalent to the mean and the interquartile range was converted to SD by dividing it by 1.35 (as per Sect. 6.5.2.5 "Interquartile ranges" of the Cochrane handbook; [67]). Discrepancies in extraction of data were solved by discussion.

Fig. 1 PRISMA Flow chart
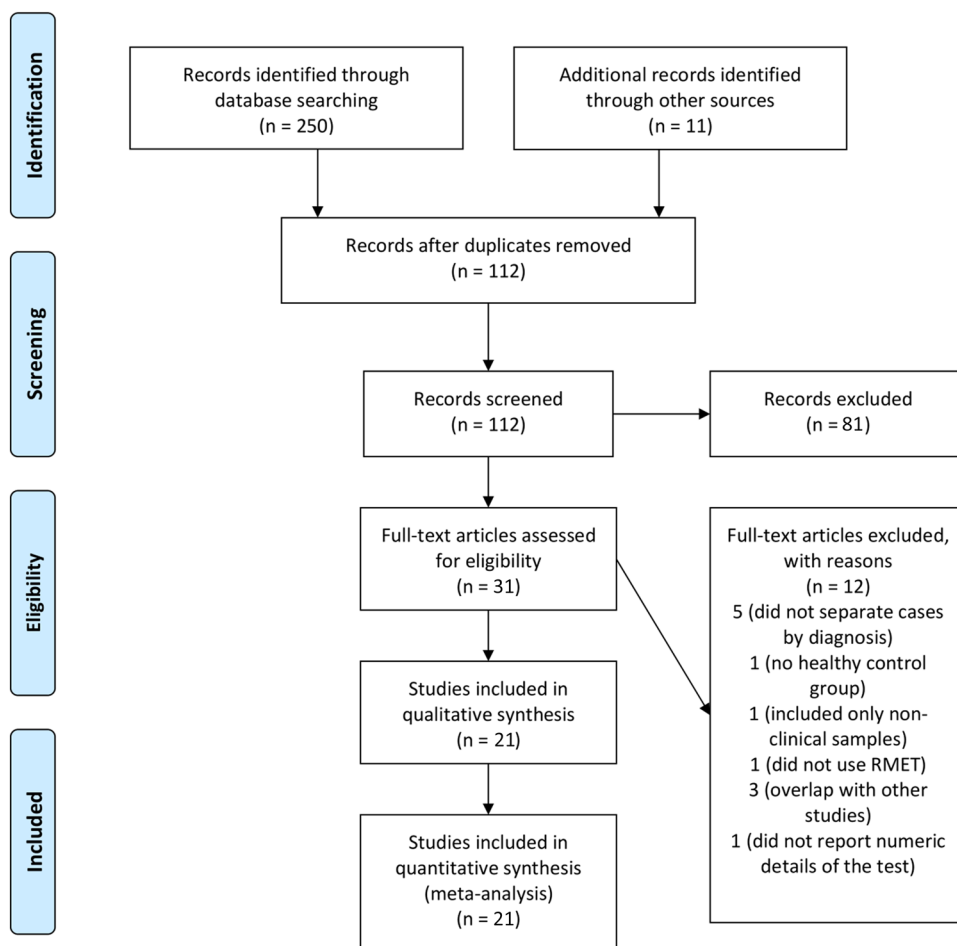
The main characteristics of the studies are summarized in Table 1; in "Results", studies are grouped by diagnosis.

The risk of bias was assessed using a validated checklist published by the US National Heart, Lung and Blood Institute for case–control studies [URL: www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools, accessed on 20 July 2021].

The procedure was implemented according to an internal protocol.

Statistical analysis

The effect size of the differences between cases and controls was calculated as bias-corrected standardized mean (i.e., Hedges's g and 95% confidence interval) and computed so that a negative value indicated an unfavorable outcome (e.g., defective performance on the RMET) [68]. According to a largely agreed rule of thumb, values of 0.20, 0.50, and 0.80 were assumed to be the thresholds for small, medium, and large effect sizes [69].

The within-study variance was estimated with the empirical Bayes estimator [70], and its 95% confidence interval (CI) was calculated using the Q-Profile method [71], with

Hartung and Knapp correction for random-effects models [72]. The Cochran's Q and I^2 statistics were used to estimate heterogeneity [73]. Significant Q statistics (i.e., $p < 0.05$) was interpreted as suggestive of heterogeneity. I^2 values between 40 and 60% were considered suggestive of moderate heterogeneity, while values above 75% were considered indicative of substantial heterogeneity [74]. The radial plot was used to assess model adequacy [75], and effect size sampling variance. For each study, the observation of a large standardized residual (above 2, as a rule of thumb) suggests that the study does not fit the assumed model (i.e., it may be an outlier). Publication bias was investigated using the funnel plot and related statistics when there were 10 or more studies [76].

We reported both the fixed-effects and the random-effects models in both tabular and graphical (forest plot) forms. Fixed-effects models estimate a common effect, which is valid only for the studies included in the meta-analysis [77]. The random-effects models aim to provide inference about the average effect in the entire population from which the studies are expected to be drawn. Essentially, the fixed-effects model assumes that variance depends principally on sampling error, while the random-effects model takes

Table 1 General characteristics of the studies

Study	Year	Location	Condition	Criteria for diagnosis	Sample	BMI	Duration of illness (years)
Harrison et al. [81]	2009	London, UK	Anorexia nervosa	DSM-IV	AN: 20 Women: 20 Age: 26.25 (5.73) HC: 20 Women: 20 Age: 28.35 (8.46)	AN: 15.81 (1.15) HC: 21.78 (1.61)	AN: 7.2 (3.2)
Russell et al. [82]	2009	London, UK	Anorexia nervosa	DSM-IV	AN: 22 Women: 22 Age: 26.7 (4.8) HC: 22 Women: 22 Age: 30.3 (6.5)	AN: 15.26 (1.2) HC: 26.2 (2.0)	9.5 (5.0)
Harrison et al. [83]	2010	London, UK	Anorexia nervosa	DSM-IV	Acute AN: 50 Women: 50 Age: 26.7 (9.82) Recovered AN: 35 Women: 35 Age: 29.0 (10.62) HC: 90 Women: 90 (?) Age: 28.5 (9.93)	Acute AN: 15.38 (1.83) Recovered AN: 21.15 (1.76) HC: 21.61 (1.89)	Acute AN: 9.23 (9.27) Recovered AN: 5.47 (2.90)
Oldershaw et al. [84]	2010	London, UK	Anorexia nervosa	DSM-IV	AN: 40 Women: 37 Age: 27.3 (10.0) Recovered AN: 27 Women: 23 Age: 29.9 (7.7) HC: 47 Women: 37 Age: 29.8 (8.0)	AN: 16.6 (1.3) Recovered AN: 20.8 (2.0) HC: 23.0 (2.8)	AN: 7.4 (8.5) Recovered AN: 5.6 3.8
Adenzato et al. [85]	2012	Brescia, Italy	Anorexia nervosa	DSM-IV	AN: 30 (AN-R: 16, AN-BP: 14) Women: 30 Age: 19.7 (6) HC: 32 Women: 32 Age: 20.5 (3)	AN: 15.0 (1.7) HC: 20.2 (1.4)	AN-R: 1.7 AN-BP: 5.8
Kenyon et al. [57]	2012	London, UK	Bulimia nervosa	DSM-IV TR	BN: 48 Women: 48 Age: 28.0 (7.7) EDNOS BN: 34 Women: 33 Age: 27.6 (6.9) HC: 57 Women: ? Age: 24.0 (2.6)	BN: 24.5 (7.1) EDNOS BN: 25.0 (6.3) HC: 22.7 (3.2)	BN: 10.7 (8.1) EDNOS BN: 9.9 (6.5)

Table 1 (continued)

Study	Year	Location	Condition	Criteria for diagnosis	Sample	BMI	Duration of illness (years)
Medina-Pradas et al. [50]	2012	Barcelona, Spain	Anorexia nervosa Bulimia nervosa EDNOS	DSM-IV TR	AN: 44 Women: 44 Age: 26.80 (5.70) BN: 30 Women: 30 Age: 26.80 (6.10) EDNOS: 23 Women: 23 Age: 26.02 (8.20) HC: 39 Women: 39 Age: 26.04 (14.70)	AN: 15.80 (1.70) BN: 20.90 (2.60) EDNOS: 25.80 (8.30) HC: 21.30 (2.10)	AN: 9.90 (6.60) BN: 10.30 (5.20) EDNOS: 7.30 (6.40)
Tapajóz Pereira de Sampaio et al. [86]	2013	Buenos Aires, Argentina	Anorexia nervosa Bulimia nervosa	DSM-IV	AN: 24 Women: 24 Age: 24.5 (7.6) BN: 24 Women: 24 Age: 24.4 (6.0) HC: 24 Women: 24 Age: 25.2 (6.9)	AN: 18.1 (1.8) BN: 25.0 (6.5) HC: 21.5 (1.8)	AN: 7.8 (5.9) BN: 7.7 (6.3)
Laghi et al. [87]	2015	Rome, Italy	Anorexia nervosa	DSM-IV TR	AN-R: 40 Women: 40 Age: 14.93 (1.48) HC: 40 Women: 40 Age: 14.88 (0.56)	AN-R: 15.76 (1.45) HC: 21.85 (2.63)	AN-R: 1.28 (1.04)
Jermakow and Brzezicka [88]	2016	Warsaw, Poland	Anorexia nervosa	ICD-10	AN: 11 Women: 11 Age: 26.8 (4.3) HC: 33 Women: 33 Age: 21.3 (1.4)	AN: 14 to 20 HC: unavailable	Unavailable
Kucharska et al. [55]	2016	Cardiff, UK	Anorexia nervosa	ICD-10 DSM-IV TR	AN: 25 Women: 25 Age: 27.1 (6.3) HC: 25 Women: 25 Age: 24.5 (5.2)	AN: 17.6 (2.2) HC: 23.4 (3.6)	AN: 9.7 (6.6)
Aloi et al. [49]	2017	Catanzaro, Italy	Binge eating disorder	DSM-5	Obese BED: 22 Women: 18 Age: 44 (11) Obese not BED: 20 Women: 9 Age: 50 (8)	BED: 36.9 Not BED: 38.2	Unavailable
Bentz et al. [89]	2017	Copenhagen, Denmark	Anorexia nervosa	ICD-10	First episode AN: 43 Women: 43 Age: 16.1 (1.5) Recovered AN: 28 Women: 28 Age: 18.4 (1.6) HC: 41 Women: 41 (?) Age: 17.7 (2.2)	First episode AN: 16.6 (1.2) Recovered AN: 21.3 (1.8) HC: 22.0 (2.6)	Unavailable

Table 1 (continued)

Study	Year	Location	Condition	Criteria for diagnosis	Sample	BMI	Duration of illness (years)
Leppanen et al. [90]	2017	London, UK	Anorexia nervosa	DSM-5	AN: 30 Women: 30 Age (median): 24.5 HC: 29 Women: 29 Age (median): 25.0	AN (median): 16.13 HC (median): 22.21	Unavailable
Redondo and Herrero-Fernández [91]	2018	Leioa, Spain	Anorexia nervosa	DSM-IV TR	AN: 38 Women: 38 Age: 21.9 (5.30) HC: 321 Women: 321 Age: 20.0 (2.05) ?	Unavailable	Unavailable
Nalbant et al. [56]	2019	Ankara, Turkey	Anorexia nervosa	DSM-5	AN: 32 Women: 32 Age: 15.2 (1.6) HC: 32 Women: 32 Age: 15.2 (1.7)	AN: 16.6 (1.5) HC: unreported	AN: 1 (0.1)
Rothschild-Yakar et al. [92]	2019	Tel Hashomer, Israel	Anorexia nervosa	DSM-5	AN: 41 (AN-R: 29 AN-BP: 12) Women: 41 Age: 17.58 (2.57) HC: 53 Women: 53 Age: 17.63 (2.39)	AN: 18.04 (1.72) HC: 21.21 (2.46)	Unavailable
Sacchetti et al. [93]	2019	London, UK	Bulimia nervosa	DSM-5	BN: 53 Women: 50 Age: 30.60 (8.91) HC: 87 Women: 76 Age: 29.14 (8.65)	BN: 23.55 (6.09) HC: 21.63 (2.72)	Unavailable
Turan et al. [94]	2019	Izmir, Turkey	Binge eating disorder	DSM-5	Obese BED: 32 Women: 18 Age: 15 (1.43) Obese not BED: 32 Women: 17 Age: 14.81 (1.65)	BED: 34.58 Not BED: 32.72	Unavailable
Konstantakopoulos et al. [58]	2020	Athens, Greece	Anorexia nervosa Bulimia nervosa	DSM-IV TR	AN: 46 Women: 46 (AN-R: 30 AN-BP: 16) Age (median): 26.0 BN: 30 Women: 30 Age (median): 24.0 HC: 42 Women: 42 Age (median): 26.0	AN (median): 16 BN (median): 19.5 HC (median): 20.7	AN (median): 5 BN (median): 4

Table 1 (continued)

Study	Year	Location	Condition	Criteria for diagnosis	Sample	BMI	Duration of illness (years)
Cortés-García et al. [95]	2021	[Unknown], [unknown]	AN	DSM-IV	AN: 44 Women: 44 Age: 15.18 (1.46) HC: 129 Women: 129 Age: 15.32 (1.17)	Unavailable	Unavailable

AN anorexia nervosa, *AN-R* AN restrictive, *AN-BP* AN binge-purge, *BN* bulimia nervosa, *BED* binge-eating disorder, *EDNOS* eating disorders not otherwise specified, *HC* healthy controls

into account heterogeneity of the studies, i.e., the fact that the effects that are estimated from the studies come from a distribution of true effects which depend on a source of variability that is not limited to sampling error [78]. In the case of 10 or more studies, we also reported an approximate prediction interval from the random-effects model for a new study, which was calculated according to Higgins et al. [79]. Prediction interval is helpful to establish whether the results of the meta-analysis will hold in future studies.

Sensitivity analysis included the application of the trim-and-fill method to ascertain the impact of publication bias on the premise that the most extreme results are not published. The trim-and-fill method recalculates the effect size by the imputation of the missing studies to produce a symmetrical funnel plot [80]. When there were enough studies ($n \geq 10$), we also recalculated the effect size by excluding outliers that were eventually revealed by the radial plot. We also used meta-regression to evaluate the impact of the following variables: year of publication; criteria for diagnosis; sample size; gender; age; body mass index; duration of illness; quality of the studies.

Meta-analysis was carried out with R (version 4.0.2) using the following packages: 'metafor' (version 2.4-0) and 'meta' (version 4.14-0).

Results

The initial search across the four bibliographic databases yielded 250 records. Additional 11 references were retrieved from past publications on the topic already known to the authors. After the exclusion of duplicates recurring in the databases, 112 articles were screened, and 31publications were inspected after excluding the articles ($n = 81$) that were judged unrelated to the topic on the basis of their title and abstract (see Fig. 1).

Twelve articles were excluded because either they had overlapping samples, or included non-clinical samples only, did not separate cases by diagnosis, did not use the RMET, or did not report numeric details (see the list of excluded studies in the supplemental material).

Overall, we retrieved 21 studies [49, 50, 53–56, 83–95]. These 21 studies included 29 independent samples of patients diagnosed with an eating disorder: 20 samples included people diagnosed with AN (17 with acute cases, and 3 with recovered cases); 5 were samples of patients diagnosed with BN; there were 2 samples of patients diagnosed with EDNOS (one specifically with EDNOS BN) and 2 samples of patients diagnosed with BED. These samples were compared with 20 independent samples of gender- and age-matched healthy controls.

Compared to a past meta-analysis on the topic [47], this systematic review and meta-analysis includes more than twice the number of studies and the sample size. Another systematic review [62], which also reported data on the RMET in eating disorders, retrieved 14 studies against 21 in the present review. A recent meta-analysis investigating autistic features in patients with anorexia nervosa and reporting data on the RMET [46], retrieved 10 independent samples only as against 17 samples in this meta-analysis. All studies included in the most recent past systematic reviews and meta-analyses on the topic are included in the present review, which also offers substantial additional information.

Table 1 lists the main characteristics of the included studies. Among these, 8 were from the UK, 3 from Italy, 2 from Spain and 2 from Turkey, and 1 each from Argentina, Denmark, Greece, Israel, and Poland. We were unable to determine the location of one study [95]. We were unable to retrieve studies from Asia, Africa, or Oceania. Thus, there is an ethnic bias in the studies about the performances on the RMET of people with eating disorders.

Sample size varied from 11 to 53 in the clinical groups (on average, 33 ± 10), and from 20 to 129 in the control groups (45 ± 25), after the exclusion of an outlier ($n = 321$ in [91]), which was, nevertheless, included in the meta-analysis.

The samples included quite exclusively girls or women (96% of all included participants).

Age varied widely in the included samples, from 15 to 44 years old in the clinical samples and from 15 to 50 years old in the control samples. Overall, there were six samples of underaged patients (< 18 years old) and just one sample

of clinical participants aged > 30 years old. We used body mass index (BMI) as a proxy of the severity of the disorder in patients with AN. As expected, BMI was lower in patients diagnosed with AN (17.2 ± 2) than in patients with BN (22.7 ± 2.4; $t = -5.07$; $df = 20$; $p < 0.0001$), and it was lower in patients with active AN (16.3 ± 1) than in those with recovered AN (21.1 ± 0.2; $t = 8.06$; $df = 15$; $p < 0.0001$). Patients with EDNOS were in the same range as those with BN (25.4 ± 0.5; $t = 1.51$; $df = 5$; $p = 0.19$), and those with BED were obese (35.7 ± 1.6). Duration of illness was reported for 20 samples and varied from 1 to 10 years.

Study quality was judged poor in 8 studies, fair in 7 studies, and good in 6 studies. Lack of sample size justification and the lack of blinding of the assessors were the most often observed shortcomings of the studies (Table A in the supplementary material).

Poor studies were equally distributed across time, while good studies were more often observed in recent years, from 2016 on.

Results of the meta-analysis

As a preliminary analysis, we estimated the raw, untransformed mean performance of the contrast group (putatively healthy controls) in the AN and BN studies. Three studies reported more than one comparison, resulting in an overlap between some of the controls included in studies on AN and those included in studies with BN [50, 58, 86]. To calculate the raw, untransformed mean performance of the controls, we excluded three studies: because they used the children version of the RMET [87, 95]; or because the authors had recalculated the global score by excluding some items from the RMET [91]. However, these studies were included in the subsequent pairwise meta-analysis since it relates to standardized mean differences, which are less likely to be affected by the version of the tool.

Mean scores on controls on the RMET were close to 27 in both the controls of AN studies (Figure A1 in supplementary material) and those of BN studies (Figure A2 in supplementary material).

Among cases, mean scores on the RMET ranged 21 to 28 (mean = 25) in patients with active AN; 24 to 28 (mean = 25) in patients with recovered AN; 24 to 27 (mean = 25) in patients with BN; 24 to 27 (mean = 25) in patients with EDNOS; and they were found to be 20 (range: 19.9 to 20.3) in the two studies including patients with BED and obesity.

Performances on the RMET of patients with AN

Overall, 17 samples were included, summarizing the results from 580 patients with active AN and 1019 controls. Ten studies found no differences between patients with active AN and controls, while seven studies found impaired performance on the RMET of patients with active AN. Overall, patients with AN in the active phase showed worse performances on the RMET than gender- and age-matched controls (details in Table 2).
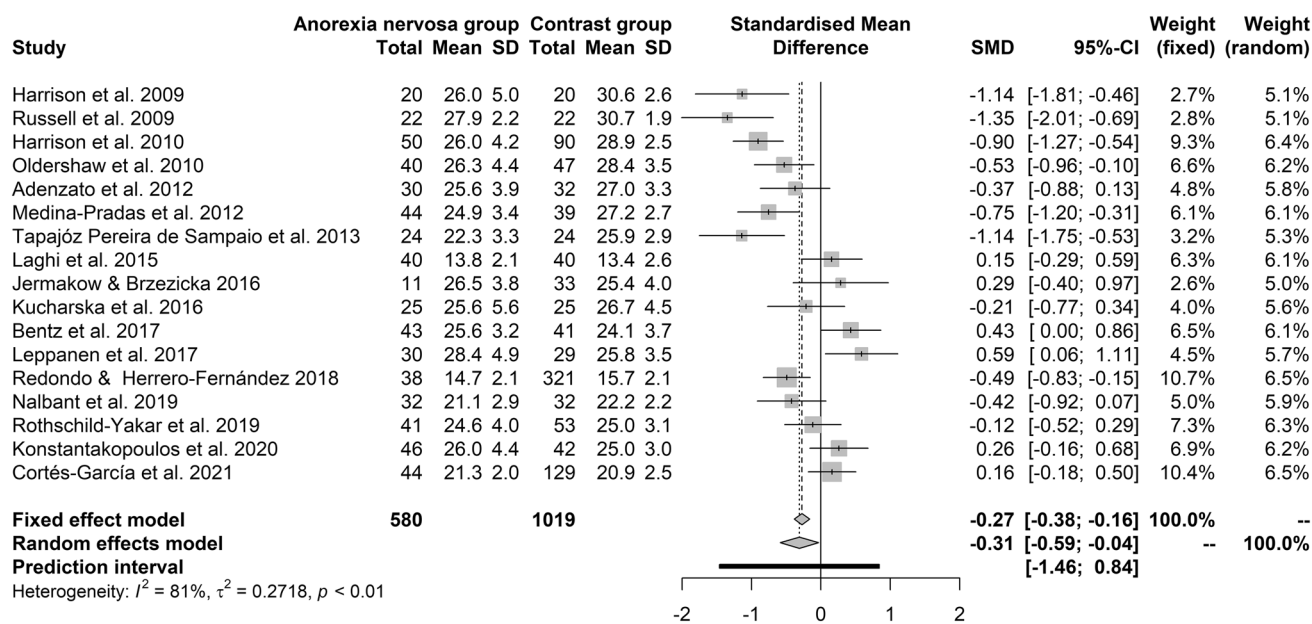
The random-effects model was more conservative than the fixed-effects model, i.e., it had a larger 95% CI, with the lower bound approaching zero; the prediction interval indicated that the finding would not hold in future studies with the same trend as those included in the analysis (Fig. 2).

Heterogeneity was substantial in this meta-analysis, likely as a reflection of differences among the samples in terms of duration of illness, type of the disorder (restrictive variant versus binge-purge variant), and stage of the condition (first-episode versus recurrence or chronicity of the episode). Earlier studies were more likely to find worse performances on the RMET of patients with active AN than later studies. Indeed, the year of publication accounted for 61% of the heterogeneity in the meta-analysis and each year imported a displacement towards better performances by 0.10 (95% CI: 0.05 to 0.15) in the differences between patients with active AN and controls (Fig. 3).

Age (Beta = -0.05; s.e. Beta = 0.03; $z = -1.84$; $p = 0.066$), BMI (Beta = -0.07; s.e. Beta = 0.08; $z = 0.40$; $p = 0.688$), and quality of the studies ($Q = 2.88$, $p = 0.236$) did not impact on estimates of the differences on the RMET between patients with active AN and controls. Conversely, criteria for diagnosis ($Q = 9.66$, $p = 0.022$) and duration of illness (Beta = -0.09; s.e. Beta = 0.04; $z = -2.49$; $p = 0.0127$) were related to the estimates and explained, respectively, 37% and 43% of heterogeneity. This was in part a consequence of their relationship with the year of publication, with longer duration of illness in samples of earlier studies (correlation between duration of illness and year of publication: $r = -0.58$, $p = 0.048$), and studies based on DSM-IV (seven studies) more likely to find differences and larger effect size (random-effects model estimates: -0.70 [-1.06 to -0.34]) than studies based on DSM-IV-TR (five studies; estimates: -0.21 [-0.62 to 0.20]), DSM-5 (three studies; estimates: 0.006 [-0.53 to 0.55]), or ICD-10 (two studies; estimates: 0.37 [-0.32 to 1.06]).

Egger's regression test did not reveal publication bias ($t = -0.759$; $df = 15$; $p = 0.46$), however, the trim-and-fill method suggested one study to be filled to gain symmetry at the funnel plot on the opposite side of the studies of Russell et al. [82] (Figure A3 in supplementary material). After reanalysis of the data with the imputed study, the random-effects model was no longer statistically significant: the estimates from 18 studies were -0.25 (95% CI: -0.54 to 0.03); $z = -1.74$; $p = 0.083$. Heterogeneity remained substantial: $I^2 = 82\%$ (95% CI: 72–88%).

The radial plot indicated that several studies scored outside two standard deviations from the mean in the sample, with the studies of Russell et al. [82] and Harrison et al.

Study	Anorexia nervosa group Total	Mean	SD	Contrast group Total	Mean	SD	Standardised Mean Difference	SMD	95%-CI	Weight (fixed)	Weight (random)
Harrison et al. 2009	20	26.0	5.0	20	30.6	2.6		-1.14	[-1.81; -0.46]	2.7%	5.1%
Russell et al. 2009	22	27.9	2.2	22	30.7	1.9		-1.35	[-2.01; -0.69]	2.8%	5.1%
Harrison et al. 2010	50	26.0	4.2	90	28.9	2.5		-0.90	[-1.27; -0.54]	9.3%	6.4%
Oldershaw et al. 2010	40	26.3	4.4	47	28.4	3.5		-0.53	[-0.96; -0.10]	6.6%	6.2%
Adenzato et al. 2012	30	25.6	3.9	32	27.0	3.3		-0.37	[-0.88; 0.13]	4.8%	5.8%
Medina-Pradas et al. 2012	44	24.9	3.4	39	27.2	2.7		-0.75	[-1.20; -0.31]	6.1%	6.1%
Tapajóz Pereira de Sampaio et al. 2013	24	22.3	3.3	24	25.9	2.9		-1.14	[-1.75; -0.53]	3.2%	5.3%
Laghi et al. 2015	40	13.8	2.1	40	13.4	2.6		0.15	[-0.29; 0.59]	6.3%	6.1%
Jermakow & Brzezicka 2016	11	26.5	3.8	33	25.4	4.0		0.29	[-0.40; 0.97]	2.6%	5.0%
Kucharska et al. 2016	25	25.6	5.6	25	26.7	4.5		-0.21	[-0.77; 0.34]	4.0%	5.6%
Bentz et al. 2017	43	25.6	3.2	41	24.1	3.7		0.43	[0.00; 0.86]	6.5%	6.1%
Leppanen et al. 2017	30	28.4	4.9	29	25.8	3.5		0.59	[0.06; 1.11]	4.5%	5.7%
Redondo & Herrero-Fernández 2018	38	14.7	2.1	321	15.7	2.1		-0.49	[-0.83; -0.15]	10.7%	6.5%
Nalbant et al. 2019	32	21.1	2.9	32	22.2	2.2		-0.42	[-0.92; 0.07]	5.0%	5.9%
Rothschild-Yakar et al. 2019	41	24.6	4.0	53	25.0	3.1		-0.12	[-0.52; 0.29]	7.3%	6.3%
Konstantakopoulos et al. 2020	46	26.0	4.4	42	25.0	3.0		0.26	[-0.16; 0.68]	6.9%	6.2%
Cortés-García et al. 2021	44	21.3	2.0	129	20.9	2.5		0.16	[-0.18; 0.50]	10.4%	6.5%
Fixed effect model	**580**			**1019**				**-0.27**	**[-0.38; -0.16]**	**100.0%**	**--**
Random effects model								**-0.31**	**[-0.59; -0.04]**	**--**	**100.0%**
Prediction interval									**[-1.46; 0.84]**		

Heterogeneity: $I^2 = 81\%$, $\tau^2 = 0.2718$, $p < 0.01$

-2 -1 0 1 2

Fig. 2 Forest plot of the effect sizes of the RMET scores' differences, calculated as Hedges' g, in the comparison between patients with active AN and controls

[83] being four standard deviations below the mean in the sample (Figure A4 in supplementary material). When the analysis was repeated by excluding these two outliers, the mean difference dropped to 0.17 in the fixed-effects model and became statistically not significant in the random-effects model (see Table 2 and Figure A5 in supplementary material).

The analysis of the three samples of patients with recovered AN confirmed a poorer performance in the patients than in controls at the fixed-effects but not at the random-effects model (Table 2 and Fig. 4).

Performances on the RMET of patients with BN

As for the patients diagnosed with BN, 5 samples were included, summarizing results from 185 patients with active BN and 249 controls. Two studies reported poorer performances of patients with BN than controls; three studies found no statistically significant differences. The meta-analysis suggested poorer performances of patients with BN than controls at the fixed-effects model, with similar effect size than in studies on patients with AN, but the random-effects model indicated that the findings did not hold when heterogeneity—which was moderate to substantial—was taken into account (Table 2 and Fig. 5, section A).

Performances on the RMET of patients with BED

Two studies explored the performances of patients with BED on the RMET and found no statistically significant

differences between 54 cases with obesity versus 52 controls with obesity but without BED (Table 2 and Fig. 5, section B).

Performances on the RMET of patients with EDNOS

Two studies included patients with EDNOS and summarized results from 57 cases, and 96 controls. One study showed poorer performances of patients than controls and the other reported no statistically significant differences. The meta-analysis of the studies including patients with EDNOS reached the statistically significant threshold on the fixed-effects model but not on the random-effects model, with moderate to substantial heterogeneity (Table 2 and Fig. 5, section C).

Discussion

At first sight, this meta-analysis found that patients with eating disorders perform worse than putatively healthy controls on a task aimed at measuring the understanding of others' mental states. However, the estimated differences had small effect sizes, which ranged from medium to negligible in the confidence interval. The differences from controls on the RMET were more clear-cut in studies including patients with active AN than in those including patients with recovered AN, or including patients with BN, and were minimal or absent in studies including patients with BED or EDNOS. The greater is the chance of malnutrition and its impact on

Table 2 Main results of the meta-analysis of studies on performances on the RMET of patients with eating disorders

		k	n	Effect size	95% CI		z	P	Q	p value	tau² (%)	I² (%)	95% CI	
Active anorexia nervosa	FE model	17	1599	− 0.27	− 0.38	− 0.16	− 4.85	0.0001						
	RE model	17	1599	− 0.31	− 0.58	− 0.04	− 2.22	0.0267	83.45	<0.001	0.272	81	70	88
	RE model after trim-and-fill	17	1643	− 0.25	− 0.54	0.03	− 1.74	0.0827	94.51	<0.001	0.322	82	72	88
	FE model without outliers	15	1415	− 0.17	− 0.29	− 0.05	− 2.86	0.0042						
	RE model without outliers	15	1415	− 0.20	− 0.47	0.06	− 1.52	0.1280	58.78	<0.001	0.212	76	61	86
Recovered anorexia nervosa	FE model	3	268	− 0.34	− 0.60	− 0.09	− 2.63	0.0086						
	RE model	3	268	− 0.31	− 0.77	0.15	− 1.30	0.1926	6.76	0.0340	0.114	70	0	91
Bulimia nervosa	FE model	5	434	− 0.38	− 0.58	− 0.19	− 3.87	0.0001						
	RE model	5	434	− 0.35	− 072	0.01	− 1.90	0.0572	12.90	0.0118	0.121	69	20	88
Binge eating disorder	FE model	2	106	− 0.12	− 0.26	0.51	− 0.64	0.52						
	RE model	2	106	− 0.12	− 0.26	0.51	− 0.64	0.52	0.27	0.6066	0.0	0	0	0
EDNOS	FE model	2	153	− 0.55	− 0.89	− 0.22	− 3.23	0.0013						
	RE model	2	153	− 0.61	− 1.26	0.04	− 1.84	0.0664	3.56	0.0592	0.159	72	0	94

k number of studies, *n* overall number of participants, *FE* fixed-effects model, *RE* random-effects model

For each diagnosis, an independent meta-analysis has been carried out. For each diagnosis, both the fixed-effects and the random-effects estimates have been reported, with sensitivity analysis when there were enough studies

brain functions [96, 97], the higher the difference between patients with eating disorders and controls. Unfortunately, studies with direct comparison by diagnosis were too few to allow meaningful analyses. Moreover, several caveats have to be taken into account to reach a full grasp of these results.

When the heterogeneity of the samples is taken into account (i.e., on the basis of the random-effects model), patients with BN, EDNOS and BED do not differ from controls. As for the patients with AN, the exclusion of the most extreme samples (the so-called outliers) makes the results of the random-effects model no longer significant. Moreover, the so-called prediction interval suggests that the findings of the meta-analysis concerning the patients with AN will not hold in future studies with the same characteristics as those included in the meta-analysis. Thus, patients with AN can show worse performances on the RMET than controls, but this depends more on some characteristics of the disorder (restrictive variant versus binge-purge variant; first-episode versus recurrent) than on their being a core symptom of the disorder itself. Studies did not offer enough details to perform subgroup analysis at the level that would be required to identify subgroups of patients with AN expressing insufficient performances on the RMET. For example, some studies found a relation between alexithymia and RMET performances, with worse performances in subjects with higher scores on a measure of alexithymia [17, 98]. Alexithymia was also suggested to affect emotion recognition [99], and emotion recognition contributes to the performances on the RMET. However, we were unable to retrieve enough studies with information on alexithymia in the investigated samples to undertake a meta-regression and evaluate the role of
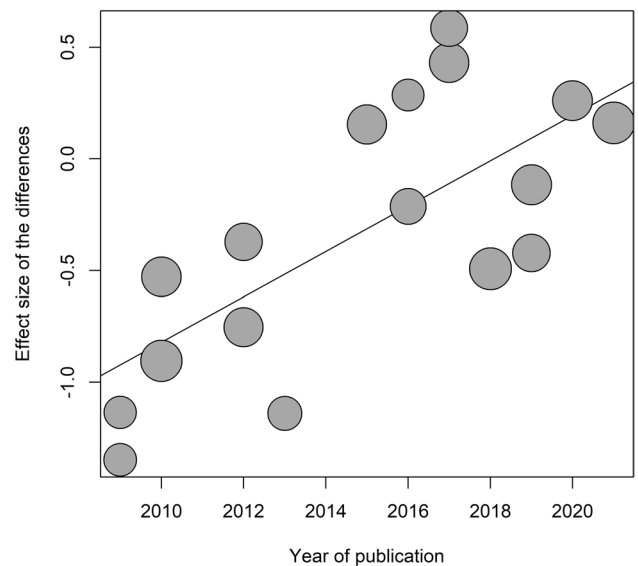
Fig. 3 Bubble plot of the result of the meta-regression of the year of publication (horizontal axis) on the effect sizes of the RMET
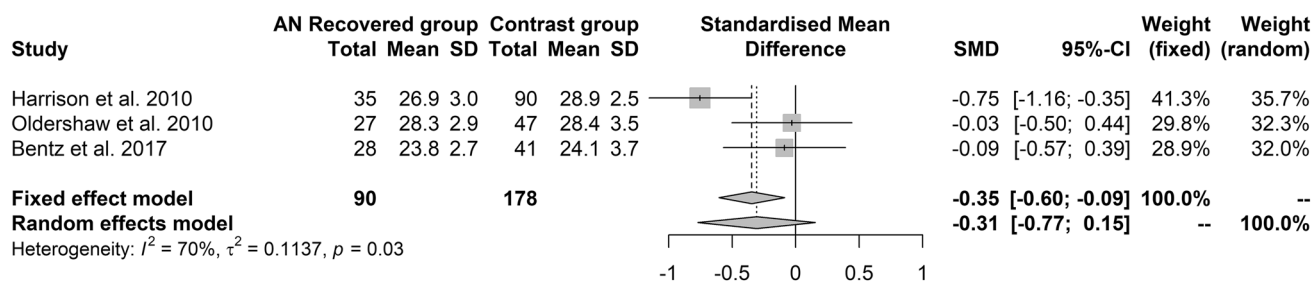
Study	AN Recovered group Total Mean SD	Contrast group Total Mean SD	Standardised Mean Difference	SMD	95%-CI	Weight (fixed)	Weight (random)
Harrison et al. 2010	35 26.9 3.0	90 28.9 2.5		-0.75	[-1.16; -0.35]	41.3%	35.7%
Oldershaw et al. 2010	27 28.3 2.9	47 28.4 3.5		-0.03	[-0.50; 0.44]	29.8%	32.3%
Bentz et al. 2017	28 23.8 2.7	41 24.1 3.7		-0.09	[-0.57; 0.39]	28.9%	32.0%
Fixed effect model	**90**	**178**		**-0.35**	**[-0.60; -0.09]**	**100.0%**	**--**
Random effects model				**-0.31**	**[-0.77; 0.15]**	**--**	**100.0%**
Heterogeneity: $I^2 = 70\%$, $\tau^2 = 0.1137$, $p = 0.03$							

-1   -0.5   0   0.5   1

Fig. 4 Forest plot of the effect sizes of the RMET scores' differences, calculated as Hedges' g, in the comparison between patients with recovered AN and controls

Study	Experimental Total Mean SD	Control Total Mean SD	Standardised Mean Difference	SMD	95%-CI
A. BN					
Kenyon et al. 2012	48 27.1 3.6	57 28.3 3.2		-0.35	[-0.74; 0.04]
Medina-Pradas et al. 2012	30 23.9 5.4	39 27.2 2.7		-0.80	[-1.30; -0.31]
Tapajóz Pereira de Sampaio et al. 2013	24 25.1 3.2	24 25.9 2.9		-0.26	[-0.83; 0.31]
Sacchetti et al. 2019	53 26.1 3.9	87 28.3 3.2		-0.63	[-0.98; -0.28]
Konstantakopoulos et al. 2020	30 26.0 3.7	42 25.0 3.0		0.30	[-0.17; 0.77]
Fixed effect model	**185**	**249**		**-0.38**	**[-0.58; -0.19]**
Random effects model				**-0.36**	**[-0.72; 0.01]**
Heterogeneity: $I^2 = 69\%$, $\tau^2 = 0.1215$, $p = 0.01$					
B. BED					
Aloi et al. 2017	22 19.9 4.0	20 18.9 3.9		0.25	[-0.36; 0.86]
Turan et al. 2019	32 20.2 2.4	32 20.1 2.2		0.04	[-0.45; 0.53]
Fixed effect model	**54**	**52**		**0.12**	**[-0.26; 0.51]**
Random effects model				**0.12**	**[-0.26; 0.51]**
Heterogeneity: $I^2 = 0\%$, $\tau^2 = 0$, $p = 0.06$					
C. EDNOS					
Kenyon et al. 2012	34 27.3 3.5	57 28.3 3.2		-0.30	[-0.73; 0.13]
Medina-Pradas et al. 2012	23 23.8 4.5	39 27.2 2.7		-0.97	[-1.51; -0.42]
Fixed effect model	**57**	**96**		**-0.55**	**[-0.89; -0.22]**
Random effects model				**-0.61**	**[-1.26; 0.04]**
Heterogeneity: $I^2 = 72\%$, $\tau^2 = 0.1594$, $p = 0.61$					

-1.5   -1   -0.5   0   0.5   1   1.5

Fig. 5 Forest plot of the effect sizes of the RMET scores' differences, calculated as Hedges' g, in the comparison between patients with BN (section **A**), BED (section **B**),

alexithymia in the patients' performances on the RMET, or its impact on the heterogeneity of the findings.

Overall, the analyses were all characterized by moderate to substantial heterogeneity (according to the calculated 95% CI of the I^2), but age, BMI, and quality of the studies did not impact estimates nor decreased the heterogeneity. The most remarkable impact on heterogeneity was from the year of the publication, which might suggest that earlier studies were published more often when they reported a difference between cases and controls, while the reverse is probably true for the most recent studies. Indeed, while Egger's test

did not reveal publication bias, the trim-and-fill method identified some asymmetry in the funnel plot. It should be taken into account that changes have occurred in the criteria for diagnosing eating disorders, and samples including patients diagnosed with AN or BN according to the DSM-5 may include less severe cases than those included in samples that were diagnosed according to the more restrictive criteria of the DSM-IV. Indeed, there is some evidence that cases that would have been classified under the label of EDNOS according to the DSM-IV are now diagnosed under one major category (AN, BN, BED) according to the DSM-5

[59]. This could be a reason why the studies using DSM-IV criteria were more likely to report that patients with AN perform worse than controls on the RMET.

Overall, this meta-analysis indicates that patients with eating disorders do not perform poorer on the RMET than gender- and age-matched healthy controls. However, patients with active AN may show some impairment on the RMET, thus revealing an impact of the disorder on the capacity of understanding others' mental state, which could affect the social functioning of the patients and might negatively influence their paths to recovery [43, 100]. There could be reasons for patients with active AN to perform poorly on a tool aimed at measuring other people's mental states. In patients with active AN, several studies based on diffusion tensor imaging reported white matter microstructural abnormalities of the left superior longitudinal and inferior frontal-occipital fasciculi, some of the bundles that connect the frontal lobe with temporal and occipital lobes [101]. The same structures were also activated during the RMET task [102]. Another region implicated in the RMET task, the inferior frontal gyrus [103], was reported to show alterations in patients with acute AN [104]. Finally, the default mode network has been consistently related to the social understanding of others' mental states [105], and has been reported to be altered in patients with acute AN [106]. Moreover, functional neuroimaging studies in patients with acute AN have shown abnormalities in the frontal visual system, the attention network, and the arousal and emotional processing systems [107], all systems contributing to the execution of the RMET task.

Some of the abnormalities reported in acute AN tend to reverse after remission of symptoms, in particular alterations in white matter microstructures [108]. Studies on patients with BN and BED reported primarily alterations in circuits involved in appetite, the processing of food stimuli, and impulse control rather than in circuits involved in the understanding of others' mental states [109], while structural and functional abnormalities are generally less extensive and pervasive in patients with BN or BED than in those with AN. Overall, brain imaging studies in patients with eating disorders are hampered by small sample sizes, inconsistencies in methods, and large variability in samples' characteristics [107–109]. Small sample sizes and large variability in samples' characteristics also affect studies on affective cognition in eating disorders [46, 62]. However, it is suggestive that abnormalities in circuits that are involved in the understanding of others' mental states are more often reported in patients with AN than in patients with BN or BED. This would be consistent with greater impairment in the RMET task in patients with AN than in patients with BN, BED or EDNOS. The likely state-dependence (because of malnutrition) of some of these abnormalities would justify better performances of patients with recovered AN than in those with acute or chronic AN.

Strength and limits

The major strengths of the study are the use of state-of-the-art methods to perform the meta-analysis and the retrieval of a larger sample of studies and participants than in the past meta-analyses on the topic. Several limitations should be also considered. Most studies had sample sizes lower than 50, and this impacted on the statistical power to find differences and apply sensitivity analyses. Participants were women in their majority, thus nothing can be said about men with eating disorders, who represent a non-negligible fraction of the cases with eating disorders, with lifetime prevalence estimated at 0.74% (95% CI: 0.24–1.52) [110]. Some relevant information—such as BMI, age of onset, or severity—was not available for all studies. Moreover, few studies were available for some specific diagnoses, thus potentially limiting the generalizability of the results. Finally, heterogeneity was high in most comparisons, and it was partially explained by year of publication, criteria for diagnosis, and duration of illness, which were interrelated to each other, limiting a clear understanding of the contribution of each element.

What is already known on this subject?

A deficit in the understanding of others' mental states, the so-called ToM, has been reported in patients with AN. Less information is available concerning other eating disorders. ToM deficits can have a role in social impairment, poor insight, and resistance to treatment that are observed in patients with eating disorders.

What does this study add?

This meta-analysis shows that patients with eating disorders do not suffer from an impaired understanding of others' mental states, with the exception of a still-to-be-identified subgroup of patients with active AN. Indeed, patients diagnosed with an active illness and who are more physically impaired because of malnutrition would exhibit poorer performance on the RMET. Publication bias might have emphasized the presence of poorer performances in patients with AN on the RMET, a recognized measure of ToM. Moreover, studies have several limitations in terms of both sample size—hence of statistical power—and data description.

Conclusions

Taking into account the results of the meta-analysis and the limitations of the included studies, it can be concluded that patients with eating disorders do not suffer from an impaired understanding of others' mental states, with the exception of a still-to-be-identified subgroup of patients with acute AN. Three main gaps should be filled in the investigation of affective cognition in eating disorders: (1) the quality of the studies should be improved, in terms of both the method and sample size; (2) more data are necessary on patients with BN, BED, or other specified feeding or eating disorder (OSFED); (3) there is a need for direct comparison between diagnoses, to confirm that the performances on the RMET are principally impaired in patients with acute AN when compared to patients with recovered AN, or with BN, BED or OSFED. Larger sample sizes and the direct comparison among diagnoses would also allow the evaluation of the role of comorbidity, especially for social anxiety and depression [111], which might affect the performances on the RMET [112, 113].

Declarations

References

1. Sebastian CL, Fontaine NM, Bird G et al (2012) Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. Soc Cogn Affect Neurosci 7(1):53–63. https://doi.org/10.1093/scan/nsr023
2. Kanske P (2018) The social mind: disentangling affective and cognitive routes to understanding others. Interdiscip Sci Rev 43:115–124. https://doi.org/10.1080/03080188.2018.1453243
3. Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? Behav Brain Sci 1:515–526
4. Sanislow CA, Ferrante M, Pacheco J et al (2019) Advancing translational research using NIMH Research domain criteria and computational methods. Neuron 101(5):779–782. https://doi.org/10.1016/j.neuron.2019.02.024
5. Baron-Cohen S, Wheelwright S, Hill J et al (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. J Child Psychol Psychiatry 42(2):241–251
6. Preti A, Vellante M, Petretto DR (2017) The psychometric properties of the "Reading the Mind in the Eyes" Test: an item response theory (IRT) analysis. Cogn Neuropsychiatry 22(3):233–253. https://doi.org/10.1080/13546805.2017.1300091
7. Black JE (2019) An IRT analysis of the reading the mind in the eyes test. J Pers Assess 101(4):425–433. https://doi.org/10.1080/00223891.2018.1447946
8. Olderbak S, Wilhelm O, Olaru G et al (2015) A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. Front Psychol 6:1503. https://doi.org/10.3389/fpsyg.2015.01503
9. Baker CA, Peterson E, Pulos S, Kirkland RA (2014) Eyes and IQ: a meta-analysis of the relationship between intelligence and 'Reading the Mind in the Eyes.' Intelligence 44:78–92
10. Vellante M, Baron-Cohen S, Melis M et al (2013) The "Reading the Mind in the Eyes" test: systematic review of psychometric properties and a validation study in Italy. Cogn Neuropsychiatry 18(4):326–354. https://doi.org/10.1080/13546805.2012.721728
11. Warrier V, Grasby KL, Uzefovsky F et al (2018) Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition. Mol Psychiatry 23(6):1402–1409. https://doi.org/10.1038/mp.2017.122

12. Adams RB Jr, Rule NO, Franklin RG Jr et al (2010) Cross-cultural reading the mind in the eyes: an fMRI investigation. J Cogn Neurosci 22(1):97–108. https://doi.org/10.1162/jocn.2009.21187

13. Castelli I, Baglio F, Blasi V et al (2010) Effects of aging on mindreading ability through the eyes: an fMRI study. Neuropsychologia 48(9):2586–2594. https://doi.org/10.1016/j.neuropsychologia.2010.05.005

14. Schiffer B, Pawliczek C, Müller BW et al (2013) Why don't men understand women? Altered neural networks for reading the language of male and female eyes. PLoS ONE 8(4):e60278. https://doi.org/10.1371/journal.pone.0060278

15. Domes G, Heinrichs M, Michel A et al (2007) Oxytocin improves "mind-reading" in humans. Biol Psychiatry 61(6):731–733. https://doi.org/10.1016/j.biopsych.2006.07.015

16. Mar RA, Oatley K, Peterson JB (2009) Exploring the link between reading fiction and empathy: ruling out individual differences and examining outcomes. Communications 34:407–428. https://doi.org/10.1515/COMM.2009.025

17. Oakley B, Brewer R, Bird G et al (2016) Theory of mind is not theory of emotion: a cautionary note on the Reading the Mind in the Eyes Test. J Abnorm Psychol 125(6):818–823. https://doi.org/10.1037/abn0000182

18. Tager-Flusberg H, Sullivan K (2000) A componential view of theory of mind: evidence from Williams syndrome. Cognition 76(1):59–90. https://doi.org/10.1016/s0010-0277(00)00069-x

19. Henry JD, von Hippel W, Molenberghs P et al (2016) Clinical assessment of social cognitive function in neurological disorders. Nat Rev Neurol 12(1):28–39. https://doi.org/10.1038/nrneurol.2015.229

20. Sato W, Kochiyama T, Uono S et al (2016) Structural neural substrates of reading the mind in the eyes. Front Hum Neurosci 10:151. https://doi.org/10.3389/fnhum.2016.00151

21. Wang X, Song Y, Zhen Z et al (2016) Functional integration of the posterior superior temporal sulcus correlates with facial expression recognition. Hum Brain Mapp 37(5):1930–1940. https://doi.org/10.1002/hbm.23145

22. Yin S, Fu C, Chen A (2018) The structural and functional correlates underlying individual heterogeneity of reading the mind in the eyes. Biol Psychol 138:179–184. https://doi.org/10.1016/j.biopsycho.2018.09.009

23. Peñuelas-Calvo I, Sareen A, Sevilla-Llewellyn-Jones J, Fernández-Berrocal P (2019) The "Reading the Mind in the Eyes" test in autism-spectrum disorders comparison with healthy controls: a systematic review and meta-analysis. J Autism Dev Disord 49(3):1048–1061. https://doi.org/10.1007/s10803-018-3814-4

24. Savla GN, Vella L, Armstrong CC et al (2013) Deficits in domains of social cognition in schizophrenia: a meta-analysis of the empirical evidence. Schizophr Bull 39(5):979–992. https://doi.org/10.1093/schbul/sbs080

25. Fernandes JM, Cajão R, Lopes R et al (2018) Social cognition in schizophrenia and autism spectrum disorders: a systematic review and meta-analysis of direct comparisons. Front Psychiatry 9:504. https://doi.org/10.3389/fpsyt.2018.00504

26. Fertuck EA, Jekal A, Song I et al (2009) Enhanced "Reading the Mind in the Eyes" in borderline personality disorder compared to healthy controls. Psychol Med 9(12):1979–1988. https://doi.org/10.1017/S003329170900600X

27. Anupama V, Bhola P, Thirthalli J, Mehta UM (2018) Pattern of social cognition deficits in individuals with borderline personality disorder. Asian J Psychiatr 33:105–112. https://doi.org/10.1016/j.ajp.2018.03.010

28. Kılıç F, Demirdaş A, Işık Ü et al (2020) Empathy, alexithymia, and theory of mind in borderline personality disorder. J Nerv Ment Dis 208(9):736–741. https://doi.org/10.1097/NMD.0000000000001196

29. Salgado RM, Pedrosa R, Bastos-Leite AJ (2020) Dysfunction of empathy and related processes in borderline personality disorder: a systematic review. Harv Rev Psychiatry 28(4):238–254. https://doi.org/10.1097/HRP.0000000000000260

30. Cusi AM, Macqueen GM, McKinnon MC (2012) Patients with bipolar disorder show impaired performance on complex tests of social cognition. Psychiatry Res 200(2–3):258–264. https://doi.org/10.1016/j.psychres.2012.06.021

31. Espinós U, Fernández-Abascal EG, Ovejero M (2018) What your eyes tell me: theory of mind in bipolar disorder. Psychiatry Res 262:536–541. https://doi.org/10.1016/j.psychres.2017.09.039

32. Dalkner N, Bengesser SA, Birner A et al (2019) The relationship between "Eyes Reading" ability and verbal memory in bipolar disorder. Psychiatry Res 273:42–51. https://doi.org/10.1016/j.psychres.2019.01.015

33. Szmulewicz A, Millett CE, Shanahan M et al (2020) Emotional processing subtypes in bipolar disorder: a cluster analysis. J Affect Disord 266:194–200. https://doi.org/10.1016/j.jad.2020.01.082

34. Zahavi AY, Sabbagh MA, Washburn D et al (2016) Serotonin and dopamine gene variation and theory of mind decoding accuracy in major depression: a preliminary investigation. PLoS ONE 11(3):e0150872. https://doi.org/10.1371/journal.pone.0150872

35. Rnic K, Sabbagh MA, Washburn D et al (2018) Childhood emotional abuse, physical abuse, and neglect are associated with theory of mind decoding accuracy in young adults with depression. Psychiatry Res 268:501–507. https://doi.org/10.1016/j.psychres.2018.07.045

36. İnanç L, Altıntaş M (2018) Are mentalizing abilities and insight related to the severity of obsessive-compulsive disorder. Psychiatry Investig 15(9):843–851. https://doi.org/10.30773/pi.2018.05.02.2

37. Yazici KU, Yazici IP (2019) Decreased theory of mind skills, increased emotion dysregulation and insight levels in adolescents diagnosed with obsessive compulsive disorder. Nord J Psychiatry 73(7):462–469. https://doi.org/10.1080/08039488.2019.1652341

38. Fett AK, Viechtbauer W, Dominguez MD et al (2011) The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 35(3):573–588. https://doi.org/10.1016/j.neubiorev.2010.07.001

39. Bora E (2017) Relationship between insight and theory of mind in schizophrenia: a meta-analysis. Schizophr Res 190:11–17. https://doi.org/10.1016/j.schres.2017.03.029

40. Merhoum N, Mengarelli F, Mottolese R et al (2015) Social functioning in autism. Key Issues Ment Health 180:46–53. https://doi.org/10.1159/000363585

41. Arcelus J, Haslam M, Farrow C, Meyer C (2013) The role of interpersonal functioning in the maintenance of eating psychopathology: a systematic review and testable model. Clin Psychol Rev 33(1):156–167. https://doi.org/10.1016/j.cpr.2012.10.009

42. Patel K, Tchanturia K, Harrison A (2016) An exploration of social functioning in young people with eating disorders: a qualitative study. PLoS ONE 11(7):e0159910. https://doi.org/10.1371/journal.pone.0159910

43. Cardi V, Mallorqui-Bague N, Albano G et al (2018) Social difficulties as risk and maintaining factors in anorexia nervosa: a mixed-method investigation. Front Psychiatry 9:12. https://doi.org/10.3389/fpsyt.2018.00012

44. Harrison A, Watterson SV, Bennett SD (2018) An experimental investigation into the use of eye-contact in social interactions in women in the acute and recovered stages of anorexia nervosa. Int J Eat Disord. https://doi.org/10.1002/eat.22993

45. Konstantakopoulos G, Ioannidi N, Patrikelis P, Gonidakis F (2020) The impact of theory of mind and neurocognition on

delusionality in anorexia nervosa. J Clin Exp Neuropsychol 42(6):611–621. https://doi.org/10.1080/13803395.2020.1786504

46. Saure E, Laasonen M, Lepistö-Paisley T et al (2020) Characteristics of autism spectrum disorders are associated with longer duration of anorexia nervosa: a systematic review and meta-analysis. Int J Eat Disord 53(7):1056–1079. https://doi.org/10.1002/eat.23259

47. 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? Int J Eat Disord 49(8):739–740. https://doi.org/10.1002/eat.22572

48. Kerr-Gaffney J, Harrison A, Tchanturia K (2018) Social anxiety in the eating disorders: a systematic review and meta-analysis. Psychol Med 48(15):2477–2491. https://doi.org/10.1017/S0033291718000752

49. Aloi M, Rania M, Caroleo M et al (2017) Social cognition and emotional functioning in patients with binge eating disorder. Eur Eat Disord Rev 25(3):172–178. https://doi.org/10.1002/erv.2504

50. Medina-Pradas C, Navarro JB, Alvarez-Moya EM et al (2012) Emotional theory of mind in eating disorders. Int J Clin Health Psychol 12:189–202

51. Albano G, Rowlands K, Baciadonna L (2019) Interpersonal difficulties in obesity: a systematic review and meta-analysis to inform a rejection sensitivity-based model. Neurosci Biobehav Rev 107:846–861. https://doi.org/10.1016/j.neubiorev.2019.09.039

52. Ivanova IV, Tasca GA, Proulx G, Bissada H (2015) Does the interpersonal model apply across eating disorder diagnostic groups? A structural equation modeling approach. Compr Psychiatry 63:80–87. https://doi.org/10.1016/j.comppsych.2015.08.009

53. McCuen-Wurst C, Ruggieri M, Allison KC (2018) Disordered eating and obesity: associations between binge-eating disorder, night-eating syndrome, and weight-related comorbidities. Ann N Y Acad Sci 1411(1):96–105. https://doi.org/10.1111/nyas.13467

54. Tonelli H, de Siqueira RL (2021) Emotion perception and theory of mind in obesity: a systematic review on the impact of social cognitive deficits on dysfunctional eating behaviors. Surg Obes Relat Dis 17(3):618–629. https://doi.org/10.1016/j.soard.2020.10.016

55. Kucharska K, Jeschke J, Mafi R (2016) Intact social cognitive processes in outpatients with anorexia nervosa: a pilot study. Ann Gen Psychiatry 15(1):24. https://doi.org/10.1186/s12991-016-0108-0

56. Nalbant K, Kalaycı BM, Akdemir D et al (2019) Emotion regulation, emotion recognition, and empathy in adolescents with anorexia nervosa. Eat Weight Disord 24(5):825–834. https://doi.org/10.1007/s40519-019-00768-8

57. Kenyon M, Samarawickrema N, Dejong H et al (2012) Theory of mind in bulimia nervosa. Int J Eat Disord 45(3):377–384. https://doi.org/10.1002/eat.20967

58. Konstantakopoulos G, Ioannidi N, Patrikelis P, Gonidakis F (2020) Empathy, mentalizing, and cognitive functioning in anorexia nervosa and bulimia nervosa. Soc Neurosci 15(4):477–488. https://doi.org/10.1080/17470919.2020.1760131

59. Call C, Walsh BT, Attia E (2013) From DSM-IV to DSM-5: changes to eating disorder diagnoses. Curr Opin Psychiatry 26(6):532–536. https://doi.org/10.1097/YCO.0b013e328365a321

60. Fernández-Sotos P, Navarro E, Torio I et al (2018) Pharmacological interventions in social cognition deficits: a systematic mapping review. Psychiatry Res 270:57–67. https://doi.org/10.1016/j.psychres.2018.09.012

61. Fernández-Sotos P, Torio I, Fernández-Caballero A et al (2019) Social cognition remediation interventions: a systematic mapping review. PLoS ONE 14(6):e0218720. https://doi.org/10.1371/journal.pone.0218720

62. Simonsen CB, Jakobsen AG, Grøntved S, Kjaersdam Telléus G (2020) The mentalization profile in patients with eating disorders: a systematic review and meta-analysis. Nord J Psychiatry 74(5):311–322. https://doi.org/10.1080/08039488.2019.1707869

63. Moher D, Liberati A, Tetzlaff J et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 62(10):1006–1012. https://doi.org/10.1016/j.jclinepi.2009.06.005

64. Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:n71. https://doi.org/10.1136/bmj.n71

65. Egger M, Jüni P, Bartlett C et al (2003) How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. Health Technol Assess 7(1):1–76

66. Ferguson CJ, Brannick MT (2012) Publication bias in psychological science: prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. Psychol Methods 17(1):120–128

67. Higgins JPT, Thomas J, Chandler J et al (2019) Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019). Cochrane. Available from www.training.cochrane.org/handbook

68. Hedges LV, Olkin I (1985) Statistical methods for meta-analysis. Academic Press, Orlando

69. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, Hillsdale, NJ

70. Veroniki AA, Jackson D, Viechtbauer W et al (2016) Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res Synth Methods 7(1):55–79. https://doi.org/10.1002/jrsm.1164

71. Viechtbauer W (2007) Confidence intervals for the amount of heterogeneity in meta-analysis. Stat Med 26(1):37–52. https://doi.org/10.1002/sim.2514

72. Knapp G, Hartung J (2003) Improved tests for a random effects meta-regression with a single covariate. Stat Med 22(17):2693–2710. https://doi.org/10.1002/sim.1482

73. Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J (2006) Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods 11(2):193–206. https://doi.org/10.1037/1082-989X.11.2.193

74. Higgins JPT, Thomas J, Chandler J et al (2021) Cochrane Handbook for systematic reviews of interventions version 6.2 (updated February 2021). Cochrane. Available from www.training.cochrane.org/handbook

75. Galbraith R (1994) Some applications of radial plots. J Am Stat Assoc 89:1232–1242

76. Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. BMJ 315(7109):629–634. https://doi.org/10.1136/bmj.315.7109.629

77. Viechtbauer W (2010) Conducting meta-analyses in R with the metaphor package. J Stat Softw 36(3):1–48

78. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2011) Introduction to meta-analysis. John Wiley & Sons, Hoboken, NJ

79. Higgins JP, Thompson SG, Spiegelhalter DJ (2009) A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc 172(1):137–159. https://doi.org/10.1111/j.1467-985X.2008.00552.x

80. Duval S, Tweedie R (2000) Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56:455–463

81. Harrison A, Sullivan S, Tchanturia K, Treasure J (2009) Emotion recognition and regulation in anorexia nervosa. Clin Psychol Psychother 16(4):348–356. https://doi.org/10.1002/cpp.628

82. Russell TA, Schmidt U, Doherty L et al (2009) Aspects of social cognition in anorexia nervosa: affective and cognitive theory of mind. Psychiatry Res 168(3):181–185. https://doi.org/10.1016/j.psychres.2008.10.028

83. Harrison A, Tchanturia K, Treasure J (2010) Attentional bias, emotion recognition, and emotion regulation in anorexia: state or trait? Biol Psychiatry 68(8):755–761. https://doi.org/10.1016/j.biopsych.2010.04.037

84. Oldershaw A, Hambrook D, Tchanturia K et al (2010) Emotional theory of mind and emotional awareness in recovered anorexia nervosa patients. Psychosom Med 72(1):73–79. https://doi.org/10.1097/PSY.0b013e3181c6c7ca

85. Adenzato M, Todisco P, Ardito RB (2012) Social cognition in anorexia nervosa: evidence of preserved theory of mind and impaired emotional functioning. PLoS ONE 7(8):e44414. https://doi.org/10.1371/journal.pone.0044414

86. Pereira T, de Sampaio F, Soneira S, Aulicino A et al (2013) Theory of mind and central coherence in eating disorders: two sides of the same coin? Psychiatry Res 210(3):1116–1122. https://doi.org/10.1016/j.psychres.2013.08.051

87. Laghi F, Pompili S, Zanna V et al (2015) Are adolescents with anorexia nervosa better at reading minds? Cogn Neuropsychiatry 20(6):489–501. https://doi.org/10.1080/13546805.2015.1091766

88. Jermakow N, Brzezicka A (2016) How autistic are anorectic females? Similarities and differences between anorexia nervosa and autism spectrum disorders. Clin Neuropsychiatry 13(4–5):53–58

89. Bentz M, Jepsen JRM, Pedersen T et al (2017) Impairment of social function in young females with recent-onset anorexia nervosa and recovered individuals. J Adolesc Health 60(1):23–32. https://doi.org/10.1016/j.jadohealth.2016.08.011

90. Leppanen J, Cardi V, Ng KW et al (2017) Effects of intranasal oxytocin on the interpretation and expression of emotions in anorexia nervosa. J Neuroendocrinol. https://doi.org/10.1111/jne.12458

91. Redondo I, Herrero-Fernández D (2018) Validation of the Reading the Mind in the Eyes Test in a healthy Spanish sample and women with anorexia nervosa. Cogn Neuropsychiatry 23(4):201–217. https://doi.org/10.1080/13546805.2018.1461618

92. Rothschild-Yakar L, Stein D, Goshen D et al (2019) Mentalizing self and other and affect regulation patterns in anorexia and depression. Front Psychol 10:2223. https://doi.org/10.3389/fpsyg.2019.02223

93. Sacchetti S, Robinson P, Bogaardt A et al (2019) Reduced mentalizing in patients with bulimia nervosa and features of borderline personality disorder: a case-control study. BMC Psychiatry 19(1):134. https://doi.org/10.1186/s12888-019-2112-9

94. Turan S, Özyurt G, Çatlı G et al (2019) Social cognition and emotion regulation may be impaired in adolescents with obesity independent of the presence of binge eating disorder: a two-center study. Psychiatr Clin Psychopharmacol 29:887–894. https://doi.org/10.1080/24750573.2019.1693727

95. Cortés-García L, Akça ÖF, Wall K, Sharp C (2021) Exploring mentalizing in adolescents with anorexia nervosa and borderline personality disorder: a comparative study of psychiatric inpatients and healthy controls. Scand J Child Adolesc Psychiatr Psychol 9:16–26. https://doi.org/10.21307/sjcapp-2021-003

96. Frank GK (2015) What causes eating disorders, and what do they cause? Biol Psychiatry 77(7):602–603. https://doi.org/10.1016/j.biopsych.2015.01.012

97. Frank GKW, Shott ME, DeGuzman MC (2019) Recent advances in understanding anorexia nervosa. F1000Res 8:504

98. Rødgaard EM, Jensen K, Mottron L (2019) An opposite pattern of cognitive performance in autistic individuals with and without alexithymia. J Abnorm Psychol 128(7):735–737. https://doi.org/10.1037/abn0000408

99. Brewer R, Cook R, Cardi V et al (2015) Emotion recognition deficits in eating disorders are explained by co-occurring alexithymia. R Soc Open Sci 2(1):140382. https://doi.org/10.1098/rsos.140382

100. Mekori E, Halevy L, Ziv SI et al (2017) Predictors of short-term outcome variables in hospitalised female adolescents with eating disorders. Int J Psychiatry Clin Pract 21(1):41–49. https://doi.org/10.1080/13651501.2016.1229794

101. Gaudio S, Carducci F, Piervincenzi C et al (2019) Altered thalamo–cortical and occipital–parietal– temporal–frontal white matter connections in patients with anorexia and bulimia nervosa: a systematic review of diffusion tensor imaging studies. J Psychiatry Neurosci 44(5):324–339. https://doi.org/10.1503/jpn.180121

102. Cabinio M, Rossetto F, Blasi V et al (2015) Mind-Reading ability and structural connectivity changes in aging. Front Psychol 6:1808. https://doi.org/10.3389/fpsyg.2015.01808

103. Thye MD, Murdaugh DL, Kana RK (2018) Brain mechanisms underlying reading the mind from eyes, voice, and actions. Neuroscience 374:172–186. https://doi.org/10.1016/j.neuroscience.2018.01.045

104. Barona M, Brown M, Clark C et al (2019) White matter alterations in anorexia nervosa: Evidence from a voxel-based meta-analysis. Neurosci Biobehav Rev 100:285–295. https://doi.org/10.1016/j.neubiorev.2019.03.002

105. Li W, Mai X, Liu C (2014) The default mode network and social understanding of others: what do brain connectivity studies tell us. Front Hum Neurosci 8:74. https://doi.org/10.3389/fnhum.2014.00074

106. Zhang S, Wang W, Su X et al (2018) Psychoradiological investigations of gray matter alterations in patients with anorexia nervosa. Transl Psychiatry 8(1):277. https://doi.org/10.1038/s41398-018-0323-3

107. Pietrini F, Castellini G, Ricca V et al (2011) Functional neuroimaging in anorexia nervosa: a clinical approach. Eur Psychiatry 26(3):176–182. https://doi.org/10.1016/j.eurpsy.2010.07.011

108. Alfano V, Mele G, Cotugno A, Longarzo M (2020) Multimodal neuroimaging in anorexia nervosa. J Neurosci Res 98(11):2178–2207. https://doi.org/10.1002/jnr.24674

109. Mele G, Alfano V, Cotugno A, Longarzo M (2020) A broad-spectrum review on multimodal neuroimaging in bulimia nervosa and binge eating disorder. Appetite 151:104712. https://doi.org/10.1016/j.appet.2020.104712

110. Qian J, Wu Y, Liu F et al (2021) An update on the prevalence of eating disorders in the general population: a systematic review and meta-analysis. Eat Weight Disord. https://doi.org/10.1007/s40519-021-01162-z

111. Marzola E, Porliod A, Panero M et al (2020) Affective temperaments and eating psychopathology in anorexia nervosa: which role for anxious and depressive traits? J Affect Disord 266:374–380. https://doi.org/10.1016/j.jad.2020.01.142

112. Alvi T, Kouros CD, Lee J et al (2020) Social anxiety is negatively associated with theory of mind and empathic accuracy. J Abnorm Psychol 129(1):108–113. https://doi.org/10.1037/abn0000493

113. Maleki G, Zabihzadeh A, Richman MJ et al (2020) Decoding and reasoning mental states in major depression and social anxiety disorder. BMC Psychiatry 20(1):463. https://doi.org/10.1186/s12888-020-02873-w