

REVIEW ARTICLE

A systematic review and meta-analysis of set-shifting ability in eating disorders

MARION E. ROBERTS^{1*}, KATE TCHANTURIA², DANIEL STAHL³,
LAURA SOUTHGATE¹ AND JANET TREASURE¹

¹ Division of Psychological Medicine, Eating Disorders Research Unit, Department of Academic Psychiatry, King's College, Guy's Hospital, London, UK; ² Division of Psychological Medicine, Eating Disorders Research Unit, Institute of Psychiatry, King's College London, UK; ³ Department of Biostatistics and Computing, Institute of Psychiatry, King's College London, UK

ABSTRACT

Background. The aim was to critically appraise and synthesize the literature relating to set-shifting ability in eating disorders. PsycINFO, Medline, and Web of Science databases were searched to December 2005. Hand searching of eating-disorder journals and relevant reference sections was also undertaken.

Method. The 15 selected studies contained both eating disorder and healthy control groups, and employed at least one of the following six neuropsychological measures of set-shifting ability; Trail Making Test (TMT), Wisconsin Card Sort Test (WCST), Brixton task, Haptic Illusion, CatBat task, or the set-shifting subset of the Cambridge Neuropsychological Test Automated Battery (CANTAB). The outcome variable was performance on the set-shifting aspect of the task. Pooled standardized mean differences (effect sizes) were calculated.

Results. TMT, WCST, CatBat and Haptic tasks had sufficient sample sizes for meta-analysis. These four tasks yielded acceptable pooled standardized effect sizes (0.36; TMT – 1.05; Haptic) with moderate variation within studies (as measured by confidence intervals). The Brixton task showed a small pooled mean difference, and displayed more variation between sample results. The effect size for CANTAB set shifting was 0.17.

Conclusion. Problems in set shifting as measured by a variety of neuropsychological tasks are present in people with eating disorders.

INTRODUCTION

Various deficits in executive functioning have been identified in the eating disorder literature, including working memory (Kemps *et al.* 2006), and response inhibition (Southgate, 2005). Set shifting, the ability move back and forth between multiple tasks, operations, or mental sets (Miyake *et al.* 2000), is a major component of

executive functioning. Problems in set shifting may manifest either as cognitive inflexibility (e.g. concrete and rigid approaches to problem solving and stimulus-bound behaviour) or response inflexibility (e.g. perseverative or stereotyped behaviours).

Three recent theoretical papers have suggested that a problem with set shifting may be a part of the risk factors for developing an eating disorder (Southgate *et al.* 2005; Tchanturia *et al.* 2005; Steinglass & Walsh, 2006) which may be linked to compulsive traits, rigidity and perfectionism (Tchanturia, 2002, 2004*a, b*). There is

* Address for correspondence: Marion E. Roberts, Division of Psychological Medicine, Eating Disorders Research Unit, Department of Academic Psychiatry, King's College, 5th Floor Thomas Guy House, Guy's Hospital, London SE1 9RT, UK.
(Email: marion.roberts@iop.kcl.ac.uk)

also suggestion that set shifting may be part of the eating disorder endophenotype, as deficits in set shifting have been found in both affected and unaffected sister pairs (Holliday *et al.* 2005).

The aim of this systematic review and meta-analysis was to collate and summarize the literature on set-shifting ability in people with eating disorders.

METHOD

The 'QUOROM statement' for meta-analyses was followed.

Searching

Papers were located using the electronic databases PsycINFO, Medline and Web of Science, by additional hand searches through reference lists and specialist eating disorder journals, and through direct contact with academic institutions with an interest in this area. Journals were searched up to December 2005. Search keyword terms were; NEUROPSYCHOLOGY, SET SHIFTING, FLEXIBILITY, RIGIDITY, MENTAL FLEXIBILITY, COGNITIVE RIGIDITY, PERSEVERATION, WISCONSIN CARD SORTING TEST, TRAIL MAKING TEST, BRIXTON, HAPTIC, CATBAT, EATING DISORDER, ANOREXIA NERVOSA, and BULIMIA NERVOSA. No date restrictions were applied to the selection of literature.

Any study employing the set-shifting tasks Trail Making Test (TMT), Wisconsin Card Sorting Test (WCST), Brixton task, Haptic Illusion, CatBat task, or the set-shifting subset of the Cambridge Neuropsychological Test Automated Battery (CANTAB) with an eating disorder population was eligible for inclusion. All selected tasks require shifting between mental sets and strategies, although the specific operations involved may differ.

The Trail Making task (Kravariti *et al.* 2003). Participants numerically connect numbered circles on a page in a 'dot-to-dot' fashion (trail A), and then alternatively link numbers and letters, i.e. 1-A-2-B-3-C (trail B). This task can be administered using pen and paper (Reitan, 1958) or, more recently, a computerized version is available (which includes an alphabetic sequence task). Time taken to complete trail B (switching task) is the measure of set-shifting ability.

Wisconsin Card Sorting Test (WCST) (Computer version 4 Psychological Corporation). Subjects are instructed to match stimulus cards with one of four category cards; a single red triangle, two green stars, three yellow crosses and four blue circles. The sorting rule changes unpredictably during the course of the task. The WCST can be administered manually by a clinician or using a computer program. The number of perseverative errors is used as a measure of set-shifting ability.

The Brixton task (Burgess & Shallice, 1997). Participants are asked to predict the movements of a blue circle on a computer screen, which changes location after each response. Occasionally, the pattern of movement changes and the participant has to abandon the old concept in favour of a new one. The total number of errors is a measure of set-shifting ability.

The Haptic Illusion task (Uznadze, 1966; Tchaturia *et al.* 2001). This is a perceptual, set-shifting task using three wooden balls. After 15 trials with balls of different sizes in their hands (with eyes closed – fixation stage), participants judge the relative size of two, same-sized balls (critical stage). The number of trials where illusions are experienced (same-sized balls perceived as different sizes) is a measure of perceptual inflexibility.

The CatBat task (Eliava, 1964; Tchaturia *et al.* 2002). Participants are asked to fill in missing letters in a written short story. In the first part of the story the context requires a 'C' (for CAT), then the context changes and 'B' (for BAT) becomes most appropriate. The number of perseverative errors ('C' where 'B' is appropriate) is the measure of set-shifting ability.

CANTAB IDED set-shifting subtest (Downes *et al.* 1989). The Cambridge intra-extra dimensional (IDED) set shift consists of stimuli (colour-filled shapes and white lines) that appear in four rectangles on a computer screen. The subject must learn the correct stimuli for selection, based on audio and visual feedback. After six correct trials (maximum 50 trials) subjects move to the next stage and the rule

shifts. Total number of errors was used as the measure of set-shifting ability.

Selection

A total of 22 studies were selected following the above search criteria. Upon inspection of the full manuscripts, three of these papers were excluded (Fox, 1981; Ferraro *et al.* 1997; Bayless *et al.* 2002), as raw data (mean and standard deviation) was not presented and was unavailable from the authors on request. A further four papers were excluded as they did not contain a healthy control group, and therefore the effect size could not be calculated (Touyz *et al.* 1986; Lauer *et al.* 2002; Kitabayashi *et al.* 2004; Frieling *et al.* 2005). A total of 15 papers are included in the systematic review. One of the selected papers was in a foreign-language journal (Koba *et al.* 2002), and another was initially in review (Steinglass *et al.* 2006), however, on submission of the current paper, this study was published.

Data abstraction

Descriptive statistics (mean, standard deviation and sample size) for eating disorder and control groups were extracted from the papers. If this data was missing it was requested from the author.

Quantitative data synthesis

Analyses were carried out in STATA 9.1 (StataCorp, College Station, TX, USA) using the user-contributed commands for meta-analyses *metan* (Bradburn *et al.* 1998) and *metabias* (Steichen, 1998).

The mean difference in scores between eating disorder and healthy control groups was standardized by calculating Cohen's *d*, the difference between the two raw means divided by the pooled standard deviation (Rosenberg *et al.* 2000). The standard error of each study's standardized effect size was calculated from the estimated effect and the group sizes of the two groups using the method of Cooper & Hedges (1994), which is implemented in *metan*.

Cohen's *d* effect sizes are defined as negligible (≥ -0.15 and > 0.15), small (≥ 0.15 and > 0.40), medium (≥ 0.40 and > 0.75), large (≥ 0.75 and > 1.10), very large (≥ 1.10 and > 1.45) and huge (≥ 1.45).

A meta-analysis was conducted for the TMT, WCST, CatBat and Haptic tasks (comparing eating disorder and healthy control groups). The four meta-analyses were conducted in the following way: The standardized effects of set-shifting ability for each task was pooled using a random-effects model, which assumes in addition to within-group variability that the mean effects differ across studies (between study heterogeneity). Random-effects models produce wider confidence intervals and are more conservative than fixed-effects models but are regarded to be more realistic due to the variety of case mix and settings (Everitt, 2003). The assumption of homogeneity of true effect sizes was assessed formally using Cochran's *Q* test of homogeneity. However, this test is not very powerful with small sample sizes and as a sample size independent measure of inconsistency I^2 was calculated [$I^2 = (Q - df)/Q$; Higgins *et al.* 2003].

Research with statistically significant results is potentially more likely to be submitted and published than studies with non-significant results. The presence of such a publication bias for the study was assessed informally by visual inspection of funnel plots [a plot of a study's precision (1/standard error) against effect size] and formally by its statistical analogue, Begg's adjusted rank test (Begg & Mazumdar, 1994), and Egger's test (Egger *et al.* 1997), which are implemented in *metabias*.

Because of a small sample size, we only present an average standardized effect size weighted by the inverse of the variance for the Brixton task.

Study characteristics

All studies employed an experimental cross-sectional design. All samples included an anorexia nervosa (AN) and healthy control population, with four studies also including bulimia nervosa (BN) patients (Jones *et al.* 1991; Tchanturia *et al.* 2001, 2004a; Murphy *et al.* 2002), and three also including AN recovered or weight-restored patients (Jones *et al.* 1991; Tchanturia *et al.* 2002, 2004b). Additionally, Murphy *et al.* (2002) included OCD patients, and Holliday *et al.* (2005) included a healthy sister comparison, however, these results will not be explored here. Little information on diagnosed co-morbidity was given, however,

Table 1. *Demographic and effect size comparison of set-shifting tasks: anorexia nervosa versus healthy control groups*

		<i>n</i>	Age (s.d.)	BMI/IBW (s.d.)	Co-morbidity	Duration of illness (years)	TMT trail B	WCST persisting errors	Brixton total	CatBat shift time	Haptic total	CANTAB shift
Steinglass <i>et al.</i> (2006)	AN	15	25.6 (6.0)	19.0 (1.0)	↑Anx ↑Dep	10.8 (5.8)	0.38	0.88				
Holliday <i>et al.</i> (2005)	HC	11	24.0 (3.1)	22.1 (1.8)		6.0 (3.0)	0.07		−0.29	0.49	0.92	
	AN	47	26.3 (10.2)	17.9 (2.7)	↑Anx ↑Dep ↑OCD							
Fowler <i>et al.</i> (2005)	HC	47	26.5 (6.1)	21.1 (2.3)		2.1 (1.4)						0.17
	AN	25	16.9 (2.0)	15.3 (1.3)	56% Dep ↑Anx							
Tchanturia <i>et al.</i> (2004 <i>b</i>)	HC	25	17.6 (2.2)	22.4 (1.3)			0.85		0.77	0.64	0.74	
	AN	34	27.2 (8.3)	13.7 (1.4)	↑Anx ↑Dep ↑OCD							
Ohrmann <i>et al.</i> (2004)	HC	36	24.9 (4.8)	21.8 (1.7)		5.5 (5.2)		0.04				
	AN	11	22.7 (3.8)	15.2 (1.2)	36% Dep							
	HC	11	27.5 (6.3)	21.3 (2.8)								
Fassino <i>et al.</i> (2002)	AN	20	23.2 (6.6)	15.6 (2.2)	n.r.			0.62				
	HC	20	23.1 (2.9)	20.6 (1.7)								
Koba <i>et al.</i> (2002)	AN	11	23.6 (5.7)	n.r.	n.r.			1.25				
	HC	7	25.9 (4.7)	n.r.								
	AN	16	22.3 (4.4)	14.8 (1.2)	↑OCD							
Murphy <i>et al.</i> (2002)	HC	16	25.3 (2.6)	22.0 (2.6)			0.01					
	AN	16	25.3 (2.6)	22.0 (2.6)								
Tchanturia <i>et al.</i> (2002)	AN	30	25.0 (6.7)	14.6 (2.1)	↑Anx ↑Dep ↑OCD				0.57		1.15	
	HC	23	27.6 (6.4)	21.3 (1.8)								
Tchanturia <i>et al.</i> (2001)	AN	15	28.1 (7.3)	14.1 (2.2)	↑Anx ↑Dep						1.63	
	HC	28	28.2 (5.6)	22.3 (2.1)								
Mathias & Kent (1998)	AN	34	22.0 (7.4)	15.3 (1.7)	↑Anx ↑Dep	3.4 (6.4)	0.44					
	HC	31	20.8 (3.6)	22.8 (2.4)								
Kingston <i>et al.</i> (1996)	AN	46	22.1 (6.7)	14.7 (1.7)	n.r.	1.7 (3.6)	0.46					
	HC	41	22.0 (5.8)	22.1 (1.9)								
	AN	10	25.8	n.r.	40% Dep ↑OCD							
Thompson (1993)	AN	10	25.8	n.r.	40% Dep ↑OCD		0.93	0.50				
	HC	10	23.2	n.r.								
Jones <i>et al.</i> (1991)	AN	30	24.4 (5.3)	59.4 (6.6)	47% Dep 7% SA	2.32 (0.8)	0.28					
	HC	39	24.9 (4.4)	98.2 (7.5)								
Witt <i>et al.</i> (1985)	AN	16	16.4 (1.9)	n.r.	n.r.		0.59					
	HC	16	16.2 (2.0)	n.r.								

AN, Anorexia nervosa; HC, healthy control; BMI, body mass index; IBW, ideal body weight; n.r., not reported; Anx, anxiety; Dep, depression; OCD, obsessive-compulsive disorder; SA, substance abuse.

Table 2. Demographic and effect size comparison of set-shifting tasks: recovered/weight-restored anorexia nervosa versus healthy control groups, and bulimia nervosa versus healthy control groups

	<i>n</i>	Age (s.d.)	BMI/IBW (s.d.)	Co- morbidity	TMT	WCST	Brixton	CATBAT	Haptic	CANTAB shift
Recovered/weight restored										
Anorexia nervosa										
Tchanturia <i>et al.</i> (2004 <i>b</i>)	ANrec	18	28.4 (6.8)	20.4 (1.5)	Nil	0.45	0.34	0.31	0.91	
	HC	36	24.9 (4.8)	21.8 (1.7)						
Tchanturia <i>et al.</i> (2002)	ANrec	16	30.0 (6.0)	20.1 (1.7)	Nil			-0.38	1.44	
	HC	23	27.6 (6.4)	21.3 (1.8)						
Jones <i>et al.</i> (1991)	ANrec	20	26.0 (6.2)	87.8 (11.2)	N.R.	0.20				
	HC	39	24.9 (4.4)	98.2 (7.5)						
Bulimia nervosa										
Tchanturia <i>et al.</i> (2004 <i>a</i>)	BN	19	26.5 (5.7)	21.8 (2.1)	↑Anx ↑Dep ↑OCD	0.52	0.07	0.94	0.88	
	HC	35	24.8 (4.7)	21.8 (1.7)						
	BN	16	22.0 (4.5)	20.1 (2.3)	↑OCD	-0.58				
Murphy <i>et al.</i> (2002)	HC	16	25.3 (2.6)	22.0 (2.6)						
	BN	15	25.1 (7.1)	20.0 (2.3)	↑Anx ↑Dep				1.40	
Tchanturia <i>et al.</i> (2001)	HC	28	28.2 (5.6)	22.3 (2.1)						
	BN	38	24.1 (4.0)	94.0 (7.3)	N.R.	0.46				
	HC	39	24.9 (4.4)	98.2 (7.5)						
Jones <i>et al.</i> (1991)	BN	38	24.1 (4.0)	94.0 (7.3)	N.R.	0.46				
	HC	39	24.9 (4.4)	98.2 (7.5)						
	HC	39	24.9 (4.4)	98.2 (7.5)						

BMI, Body mass index; IBW, ideal body weight; ANrec, Anorexia nervosa recovered; HC, healthy control; NR, Not reported; Anx, anxiety; Dep, depression; OCD, obsessive-compulsive disorder; SA, substance abuse.

a number of studies reported histories of substance abuse (e.g. Jones *et al.* 1991) and depression (Jones *et al.* 1991; Thompson, 1993; Ohrmann *et al.* 2004; Fowler *et al.* 2005). (See Tables 1 and 2 for further information regarding co-morbidity.)

The case mix studied showed wide variation: age, body mass index (BMI), diagnosis and duration of illness were noted for each sample, in order to assess clinical heterogeneity. 'Recovered AN' were classified as those who had maintained a stable BMI of 19–24 for a minimum of 1 year (Tchanturia *et al.* 2002, 2004*b*). 'Weight restored AN' were classified as those who had maintained weight for a minimum of 6 months (mean = 47.1 months, s.d. = 31.6; Jones *et al.* 1991). 'Broad AN' were classified as those groups where not all participants fulfilled the criteria for AN on weight. It was not possible to note how many cases of AN had previous episodes of BN, or vice versa (an exception being Koba *et al.* 2002, where 50% of participants also had BN symptoms). Results from these subpopulations were kept separate in the analysis. (See Tables 1 and 2 for age, BMI, duration of illness, and co-morbidity information for each sample.)

RESULTS

Trail Making test (TMT)

The TMT was the most commonly employed measure (Witt *et al.* 1985; Jones *et al.* 1991; Thompson, 1993; Kingston *et al.* 1996; Mathias & Kent, 1998; Murphy *et al.* 2002; Tchanturia *et al.* 2004*a,b*; Holliday *et al.* 2005; Steinglass *et al.* 2006). A meta-analysis of TMT shifting performance revealed a small pooled standardized mean difference of 0.36 (see Fig. 1). There was no evidence of heterogeneity [$\chi^2(13) = 16.68$, $p = 0.21$] between the studies, i.e. between AN and BN and with different states of severity and recovery. The effect sizes across studies were found to be consistent ($I^2 = 0.11$). Begg's funnel plot suggests that little publication bias was present (see Fig. A2 in online Appendix) and both Begg's and Egger's tests for publication bias were non-significant ($p = 0.91$, $p = 0.97$ respectively). Analysis for correction of publication bias (trim-and-fill method) revealed little difference in the results, therefore uncorrected data is presented here.

Wisconsin Card Sort Test (WCST)

Five papers (Thompson, 1993; Fassino *et al.* 2002; Koba *et al.* 2002; Ohrmann *et al.* 2004;

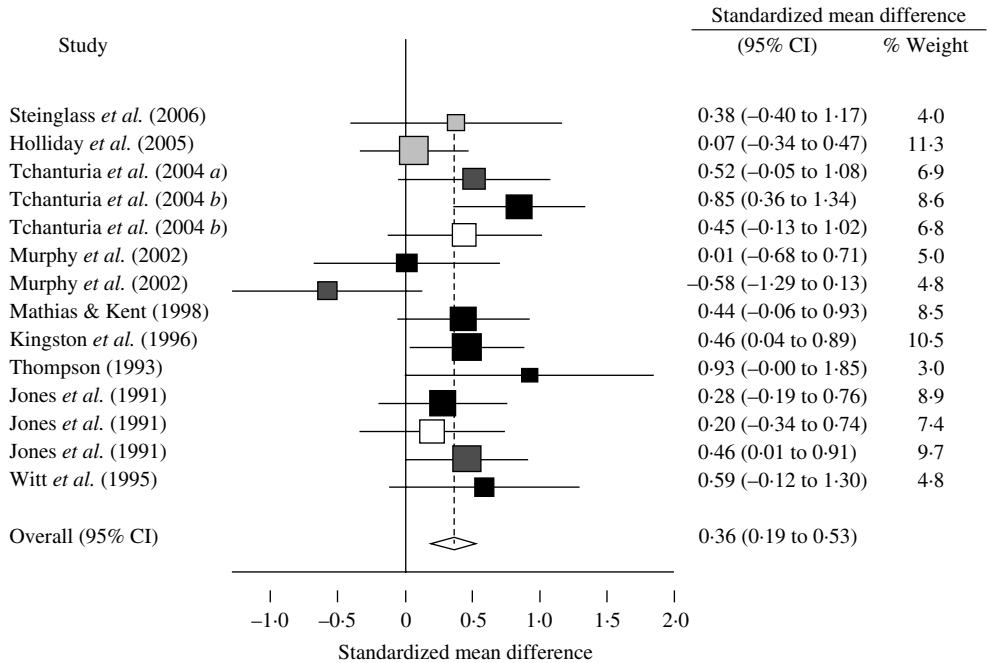


FIG. 1. Forrest plot for TMT meta-analysis. ■, Anorexia nervosa (AN); □, AN Recovered; ▣, AN Broad; ▤, Bulimia nervosa.

Steinglass *et al.* 2006) employed the WCST with an AN population. The meta-analysis of WCST persisting errors produced a pooled standardized mean difference of 0.62 (see Fig. 2). There was no evidence for heterogeneity [$\chi^2(4)=3.73$, $p=0.44$] (including one study with broad criteria) or publication bias (Begg's $p=0.81$; Egger's $p=0.64$) (see Fig. A3 in the online Appendix), and effect size was consistent across studies ($I^2=-0.07$).

The only paper employing the WCST with a BN population was not included in the meta-analysis as raw data was unavailable (Ferraro *et al.* 1997), however, the authors noted a significant deficit in BN compared to healthy control performance on this task. The BN group also displayed significantly more variance in their scores than controls.

Brixton task

Three published studies employed the Brixton task (Tchanturia *et al.* 2004a,b; Holliday *et al.* 2005), all of them from our research group. No meta-analysis was calculated for the Brixton task, as there were only four data-points

across studies and eating disorder groups. An average standardized effect size of 0.21 was calculated for the Brixton task. It should be noted that wide variation in effect size was noted across samples employing this task (see Tables 1 and 2). The only group in which the confidence interval did not overlap with zero were people acutely ill with AN (Tchanturia *et al.* 2004b).

Haptic Illusion

The Haptic Illusion is another measure that has only been used by our research group. A meta-analysis of errors on this task yielded a large pooled standardized mean difference of 1.05 (see Fig. 3). There was no evidence for heterogeneity [$\chi^2(7)=7.11$, $p=0.42$] between BN or AN samples, or in AN samples with broad criteria or weight recovery. Evidence of publication bias was found (Begg's $p=0.03$, Egger's $p=0.01$), however, this is within the 95% confidence interval limits (see Fig. A3 in the online Appendix), and the trim-and-fill method did not predict any change in the data. Also, given the large overall effect size, it can be concluded that

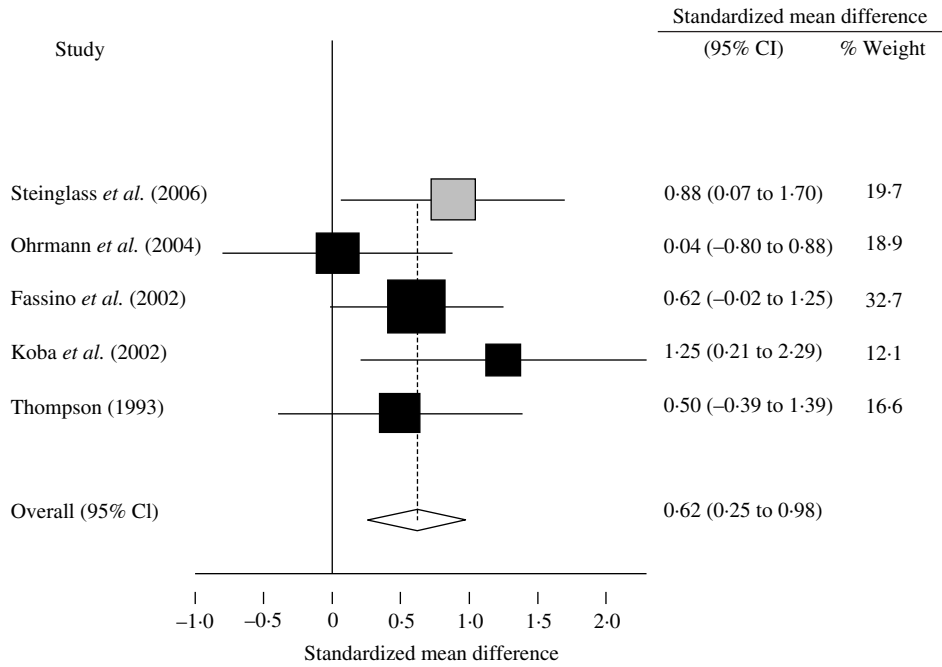


FIG. 2. Forrest plot for WCST meta-analysis. ■, Anorexia nervosa (AN); □, AN Recovered; ▨, AN Broad; ▩, Bulimia nervosa.

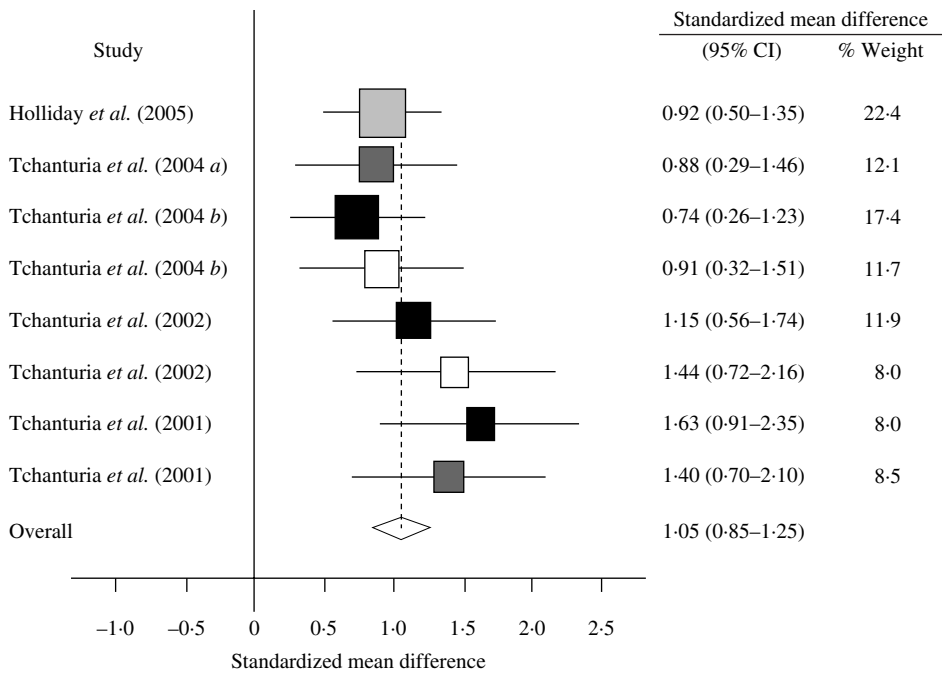


FIG. 3. Forrest plot for Haptic task meta-analysis. ■, Anorexia nervosa (AN); □, AN Recovered; ▨, AN Broad; ▩, Bulimia nervosa.

this finding is reliable. Effect size was consistent across studies ($I^2 = 0.016$).

CatBat task

The CatBat task is the third measure that has only been used by our research group (Tchanturia *et al.* 2002, 2004*a, b*; Holliday *et al.* 2005). A meta-analysis of CatBat performance revealed a medium pooled standardized mean difference of 0.45 (see Fig. A1 in the online Appendix). Heterogeneity was non-significant. No evidence for publication bias was found (Begg's $p = 0.71$, Egger's $p = 0.47$).

CANTAB set shifting

Only one study was found that employed the CANTAB IDED shifting subtest in an eating disorder population (Fowler *et al.* 2005). No difference was found between 25 AN and 25 healthy control participants, with a small effect size of 0.17.

DISCUSSION

This paper reviewed 15 studies that administered at least one of six neuropsychological set-shifting tasks in eating disorder populations. A consistent deficit in set-shifting ability was found across diagnoses, state of illness and most of the set-shifting measures. It was possible to conduct a meta-analysis of studies for TMT, WCST, Haptic, and CatBat tasks. The size of the pooled effect size varied between tasks, from small (TMT B), to medium (WCST and CatBat task), to large (Haptic task). The Brixton task has been less widely used and may only show an effect in the acute state. The set-shifting subtest of the CANTAB was used once and had an effect size close to negligible.

The limited amount of data from the recovered/weight-restored subgroups of AN suggested that the deficit in set shifting in some tests (TMT, Haptic, CatBat) remains as a trait and might be a candidate endophenotype. Further research is required to investigate this possibility. Likewise, the available data from people with BN was restricted, but suggests that the deficit in set shifting measured with TMT, CatBat task, and Haptic task is similar to that of AN whereas the Brixton task showed no effect.

There appears to be a difference in the potency of the set-shifting tasks that have been used. The Haptic task clearly has the highest, most consistent effect sizes. It is interesting to note that was the only task employed where set shifting was measured perceptually. Grunwald *et al.* (2001*a, b*) assessed Haptic performance in AN (before and after weight recovery) and control women by asking them to reproduce a tactile pattern they had traced with their fingers. The drawings of AN women were of considerably poorer quality regardless of state of illness, suggesting a general deficit in somatosensory or Haptic processing. It is possible that as these Haptic tasks are initially more difficult, administering a Haptic task as a measure of set shifting served to magnify effect size, thereby producing the larger effect sizes seen for the Haptic task in this paper.

From the limited data available the deficit in set shifting is found across eating disorder diagnoses. Interestingly this deficit is also found in other psychiatric conditions. We searched for meta-analysis/systematic review/review and ADHD, psychiatric, manic depressive, psychosis, etc. and found that a set-shifting deficit is not specific to the eating disorder population. In a systematic review of cognitive deficits including set shifting in euthymic patients with bipolar disorder, effect sizes were between 0.5 and 0.8 (Robinson *et al.* 2006) and in adult ADHD effect size was 0.65. Set-shifting abnormalities have also been found in the first-degree relatives of people with bipolar I disorder (Clark *et al.* 2005) and with schizophrenia (Snitz *et al.* 2006). Effect size for the WCST was small for people with OCD (Henry, 2006), who display a larger effect size on the CANTAB IDED shift (Chamberlain *et al.* 2006). Thus, it appears that weak set shifting is an endophenotype that broadly increases the risk of many forms of psychiatric illnesses.

A number of limitations in the current literature have been identified throughout this review. First, the majority of studies employ small samples. Second, we excluded longitudinal datasets as, to our knowledge, there is no evidence that these tasks are reliable if used repeatedly. Third, the Haptic, Brixton, and CatBat tasks have been employed only by our research group to date. Replication by other research groups among differing samples is

required in order to validate the findings presented here. Finally, it is unfortunate that only one study employed the CANTAB IDED shifting task. An explanation for the small effect size of this task could be found in the case mix of the Fowler *et al.* (2005) study. This population differs markedly from the other populations presented here, as it is an adolescent group with a short duration of illness. This is relevant because the diagnosis of AN is unstable in the early phase, as many cases recover or evolve into BN. These uncertainties exemplify the need for further work in this area.

Nevertheless this research underpins new forms of treatment in which the set-shifting deficit is a candidate trait for cognitive remediation therapy. Training the patient to adopt a more flexible thinking style in approaching everyday tasks could also encourage flexible thinking around eating disorder pathology. Such an intervention is currently being trialled in our group, and case studies thus far are promising (Davies & Tchanturia, 2005; Tchanturia *et al.* in press).

This review highlights a consistent set-shifting deficit in the eating disorder population. Many interesting questions remain. Are these traits linked to Axis I and II co-morbidity such as obsessive-compulsive traits? Are these traits also present in first-degree relatives of those with eating disorders and might they be an endophenotype? Are these traits linked to a specific genotype? To what degree do these traits become exaggerated in the acute phase of the illness? How do these general risk factors for psychopathology interact with other variables to produce eating disorder psychopathology? Do these traits affect outcome? Can these traits be a useful focus for treatment? We suggest that research examining this potential endophenotype or intermediate phenotype may lead to interesting new developments in the field.

ACKNOWLEDGEMENTS

We thank David Roberts for his assistance with figure design, and Michiko Nakazato for her assistance with foreign correspondence.

DECLARATION OF INTEREST

None.

NOTE

Supplementary information accompanies this paper on the Journal's website (<http://journals.cambridge.org>).

REFERENCES

- Bayless, J. D., Kanz, J. E., Moser, D. J., McDowell, B. D., Bowers, W. A., Andersen, A. & Paulson, J. S. (2002). Neuropsychological characteristics of patients in a hospital-based eating disorder program. *Annals of Clinical Psychiatry* **14**, 203–207.
- Begg, C. B. & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics* **50**, 1088–1101.
- Bradburn, M. J., Deeks, J. J. & Altman, D. G. (1998). *sbe24: metan – an alternative meta-analysis command*. *Stata Technical Bulletin* **44**, 1–15.
- Burgess, P. W. & Shallice, T. (1997). *The Hayling and Brixton Tests*. Thames Valley Test Company Ltd: Bury St Edmunds, UK.
- Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Robbins, T. W. & Sahakian, B. J. (2006). Motor inhibition and cognitive flexibility in obsessive-compulsive disorder and trichotillomania. *American Journal of Psychiatry* **163**, 1282–1284.
- Clark, L., Sarna, A. & Goodwin, G. M. (2005). Impairment of executive function but not memory in first-degree relatives of patients with bipolar I disorder and in euthymic patients with unipolar depression. *American Journal of Psychiatry* **162**, 1980–1982.
- Cooper, H. & Hedges, L. (1994). *The Handbook of Research Synthesis*. The Russell Sage Foundation: New York.
- Davies, H. & Tchanturia, K. (2005). Cognitive remediation therapy as an intervention for acute anorexia nervosa: a case report. *European Eating Disorders Review* **13**, 311–316.
- Downes, J. J., Roberts, A. C., Sahakian, B. J., Evenden, J. L., Morris, R. G. & Robbins, T. W. (1989). Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: evidence for a specific attentional dysfunction. *Neuropsychologia* **27**, 1329–1343.
- Egger, M., Smith, G. D., Schneider, M. & Minder, C. (1997). Bias in meta analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.
- Eliava, N. (1964). *A Problem of Set in Cognitive Psychology*. Academic Press: Tbilisi, Georgia.
- Everitt, B. S. (2003). *Modern Medical Statistics*. Arnold Publishers: London.
- Fassino, S., Piero, A., Daga, G. A., Leombruni, P., Mortara, P. & Rovers, G. G. (2002). Attentional biases and frontal functioning in anorexia nervosa. *International Journal of Eating Disorders* **31**, 274–283.
- Ferraro, F. R., Wonderlich, S. & Jolic, Z. (1997). Performance variability as a newtheoretical mechanism regarding eating disorders and cognitive processing. *International Journal of Eating Disorders* **53**, 117–121.
- Fowler, L., Blackwell, A., Jaffa, A., Palmer, R., Robbins, T. W., Sahakian, B. J. & Dowson, J. H. (2005). Profile of neurocognitive impairments associated with female in-patients with anorexia nervosa. *Psychological Medicine* **36**, 517–527.
- Fox, C. F. (1981). Neuropsychological correlates of anorexia nervosa. *International Journal of Psychiatry in Medicine* **11**, 285–290.
- Frieling, H., Roschke, B., Kornhuber, J., Wilhelm, J., Romer, K. D., Gruss, B., Donsch, D., Hillemacher, T., de Zwaan, M., Jacoby, G. E. & Bleich, S. (2005). Cognitive impairment and its association with homocysteine plasma levels in females with eating disorders – findings from the HEAd study. *Journal of Neural Transmission* **112**, 1591–1598.
- Grunwald, M., Etrich, C., Assmann, B., Dahne, A., Krause, W., Busse, F. & Gertz, H.-J. (2001a). Deficits in haptic perception

- and right parietal theta power changes in patients with anorexia nervosa before and after weight gain. *International Journal of Eating Disorders* **29**, 417–428.
- Grunwald, M., Ettrich, C., Krause, W., Assmann, B., Dahne, A., Weiss, T. & Gertz, H.-J. (2001b). Haptic perception in anorexia nervosa before and after weight gain. *Journal of Clinical and Experimental Neuropsychology* **23**, 520–529.
- Henry, J. D. (2006). A meta-analytic review of Wisconsin card sorting test and verbal fluency performance in obsessive-compulsive disorder. *Cognitive Neuropsychiatry* **11**, 156–176.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal* **327**, 557–560.
- Holliday, J., Tchanturia, K., Landau, S., Collier, D. & Treasure, J. (2005). Is impaired set shifting an endophenotype of anorexia nervosa? *American Journal of Psychiatry* **162**, 2269–2275.
- Jones, B. P., Duncan, C. C., Brouwers, P. & Mirsky, A. F. (1991). Cognition in eating disorders. *Journal of Clinical and Experimental Neuropsychology* **13**, 711–728.
- Kemps, E., Tiggemann, M., Wade, T., Ben-Tovim, D. & Breyer, R. (2006). Selective working memory deficits in anorexia nervosa. *European Eating Disorders Review* **14**, 97–103.
- Kingston, J., Szmukler, G., Andrews, D., Tress, B. & Desmond, P. (1996). Neuropsychological and structural brain changes in anorexia nervosa before and after refeeding. *Psychological Medicine* **26**, 15–28.
- Kitabayashi, Y., Ueda, H., Kashima, A., Okamoto, A., Kooguchi, Y., Narumoto, J., Wada, Y., Yamashita, T. & Fukui, K. (2004). Change of cognitive function following body weight recovery in anorexia nervosa. *Seishin Igaku* **46**, 243–248.
- Koba, T., Shrie, S. & Nabeta, Y. (2002). Impaired performance of Wisconsin Card Sorting Test in patients with eating disorders: a preliminary study. *Seishin Igaku* **44**, 681–683.
- Kravariti, E., Morris, R. G., Rabe-Hesketh, D., Murray, R. M. & Frangou, S. (2003). The Maudsley early onset schizophrenia study: cognitive function in adolescents with recent onset schizophrenia. *Schizophrenia Research* **61**, 137–148.
- Lauer, C. J., Gorzewski, B., Gerlinghoff, M., Backmund, H. & Zihl, J. (2002). Neuropsychological assessments before and after treatment in patients with anorexia nervosa and bulimia nervosa. *Journal of Psychiatric Research* **33**, 129–138.
- Mathias, J. L. & Kent, P. S. (1998). Neuropsychological consequences of extreme weight loss and dietary restriction in patients with anorexia nervosa. *Journal of Clinical and Experimental Neuropsychology* **20**, 548–564.
- Miyake, A., Freidman, N. P., Emerson, M. J., Wizki, A. H., Howter, A. & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cognitive Psychology* **41**, 49–100.
- Murphy, R., Nutzinger, D. O., Paul, T. & Leplow, B. (2002). Dissociated conditional-associative learning in anorexia nervosa. *Journal of Clinical and Experimental Neuropsychology* **24**, 176–186.
- Ohrmann, P., Kersting, A., Suslow, T., Lalee-Mentzel, J., Donges, U. S., Feibich, M., Arolt, V., Heindel, W. & Pfleiderer, B. (2004). Proton magnetic resonance spectroscopy in anorexia nervosa: Correlations with cognition. *Neuroreport* **15**, 549–553.
- Reitan, R. M. (1958). Validity of the trail making test as indicator of organic brain damage. *Perceptual and Motor Skills* **8**, 271–276.
- Robinson, L. J., Thompson, J. M., Gallagher, P., Goswami, U., Young, A. H., Ferrier, I. N. & Moore, P. B. (2006). A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *Journal of Affective Disorders* **93**, 105–115.
- Rosenberg, M. S., Adams, D. C. & Gurevitch, J. (2000). *MetaWin: Statistical Software for Meta-analysis*. Version 2.0. Sinauer: Sunderland, MA.
- Snitz, B. E., Macdonald III, A. W. & Carter, C. S. (2006). Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. *Schizophrenia Bulletin* **32**, 179–194.
- Southgate, L. (2005). Response inhibition in anorexia nervosa and bulimia nervosa: an exploration of neuropsychological functions and their association with personality traits and behaviours. Ph.D. thesis. Kings College London, Institute of Psychiatry.
- Southgate, L., Tchanturia, K. & Treasure, J. (2005). Building a model of the aetiology of eating disorders by translating experimental neuroscience into clinical practice. *Journal of Mental Health* **14**, 553–566.
- Steichen, T. (1998). sbe19: tests for publication bias in meta-analysis. *Stata Technical Bulletin* **41**, 9–15.
- Steinglass, J. & Walsh, B. T. (2006). Habit learning and anorexia nervosa: a cognitive neuroscience hypothesis. *International Journal of Eating Disorders* **39**, 267–275.
- Steinglass, J. E., Walsh, B. T. & Stern, Y. (2006). Set shifting deficit in anorexia nervosa. *Journal of the International Neuropsychological Society* **12**, 431–435.
- Tchanturia, K., Brecelj Anderlug, M., Morris, R. G., Rabe-Hesketh, S., Collier, D., Sanchez, P. & Treasure, J. L. (2004a). Cognitive Flexibility in anorexia nervosa and bulimia nervosa. *Journal of the International Neuropsychological Society* **10**, 513–520.
- Tchanturia, K., Campbell, I., Morris, R. & Treasure, J. (2005). Neuropsychological studies in anorexia nervosa. *International Journal of Eating Disorders* **37** (Suppl.), S72–S76.
- Tchanturia, K., Morris, R. G., Brecelj Anderlug, M., Collier, D. A., Nikolaou, V. & Treasure, J. (2004b). Set shifting in anorexia nervosa: an examination before and after weight gain, in full recovery, and relationship to childhood and adult OCD traits. *Journal of Psychiatric Research* **38**, 545–552.
- Tchanturia, K., Morris, R., Surguladze, S. & Treasure, J. (2002). An examination of perceptual and cognitive set shifting tasks in acute anorexia nervosa and following recovery. *Eating and Weight Disorders* **7**, 312–316.
- Tchanturia, K., Serpell, L., Troop, N. & Treasure, J. (2001). Perceptual illusions in eating disorders: rigid and fluctuating styles. *Journal of Behaviour Therapy and Experimental Psychiatry* **32**, 107–115.
- Tchanturia, K., Whitney, J. & Treasure, J. (in press). Can cognitive exercises help treat anorexia nervosa? A Case Report. *Weight and Eating Disorders* **11**.
- Thompson, S. B. N. (1993). Implications of neuropsychological test results of women in a new phase of anorexia nervosa. *European Eating Disorders Review* **1**, 152–165.
- Touyz, S. W., Beumont, P. J. V. & Johnstone, L. C. (1986). Neuropsychological correlates of dieting disorders. *International Journal of Eating Disorders* **5**, 1025–1034.
- Uznadze, D. N. (1966). *The Psychology of Set*. Consultants' Bureau: New York.
- Witt, E. D., Ryan, C. & George Hsu, L. K. (1985). Learning deficits in adolescents with anorexia nervosa. *Journal of Nervous and Mental Disease* **173**, 182–184.

APPENDIX

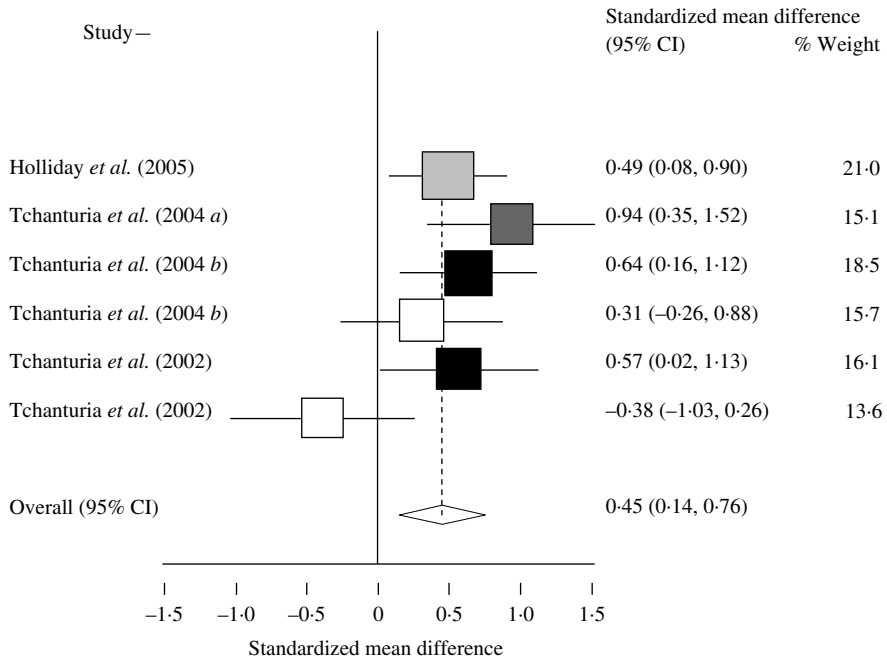


FIG. A1. Forrest plot for CatBat Task meta-analysis.

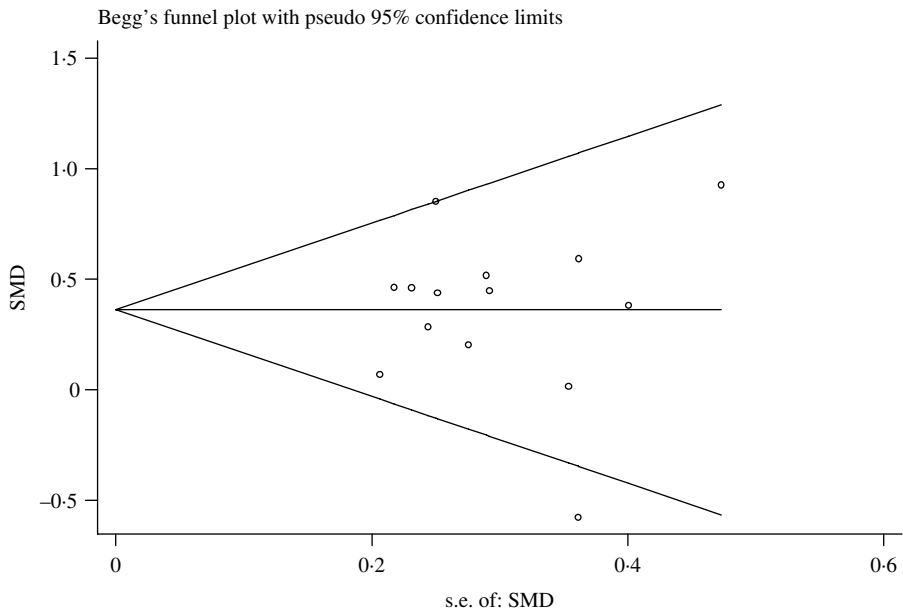


FIG. A2. Begg's funnel plot (assessing publication bias) for TMT meta-analysis.

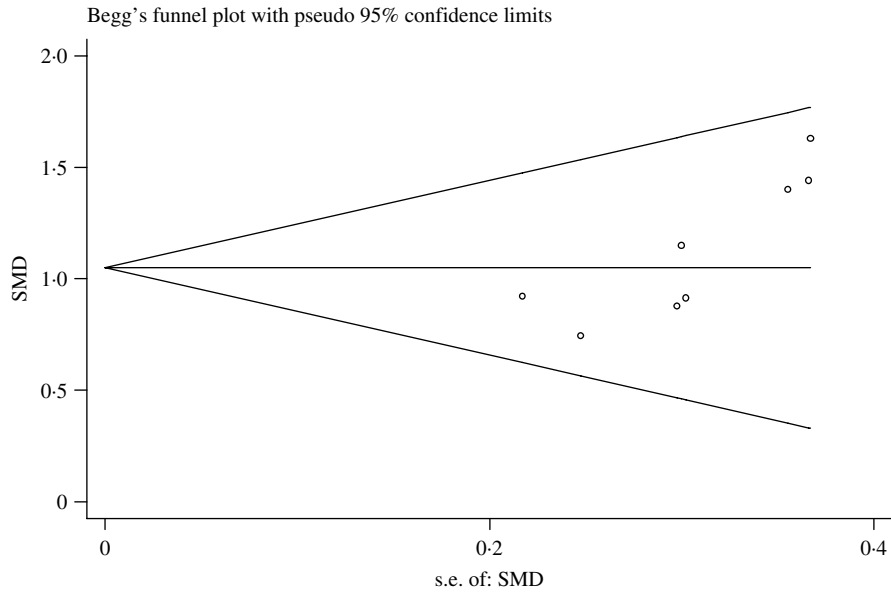


FIG. A3. Begg's funnel plot (assessing publication bias) for Haptic meta-analysis.