An Investigation into Neuropsychological Profiles in Anorexia Nervosa and Associated Clinical and Demographic Variables

Submitted by Mr Alexander Charles Lightley Drake, to the University of Exeter

as a thesis for the degree of Doctor of Clinical Psychology, 2nd May 2019.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature:…….

Contents

Literature Review

Empirical Paper

Tables

Literature Review

Empirical Paper

Figures

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

LITERATURE REVIEW

A Systematic Literature Review of Neuropsychological Findings in Anorexia Nervosa Subtypes and the Neuropsychological Profiles of Anorexia Nervosa

Trainee Name:	**Mr Alexander Charles Lightley Drake**
Primary Research Supervisor:	**Dr Ian Frampton**
	Consultant Clinical Psychologist
Secondary Research Supervisor:	**Dr Anke Karl**
	Senior Lecturer
Target Journal:	European Eating Disorders Review
Word Count:	5998 words (excluding abstract, table of contents, list of figures, references, footnotes, appendices)

Submitted in partial fulfilment of requirements for the Doctorate Degree in Clinical Psychology, University of Exeter

Abstract

Anorexia nervosa (AN) continues to be poorly understood. Substantial research has investigated the concomitant neuropsychological features of the disorder, though considerable inconsistencies remain. AN subtypes have long been described but little is known about their neuropsychological distinctions. Recent authors have suggested that there may exist within AN several neuropsychological profiles. A systematic literature review was conducted to establish whether AN neuropsychological performance could be used to identify distinct neuropsychological clusters within patients and what evidence existed that AN diagnostic subtypes displayed different neuropsychological characteristics. Furthermore, evidence was sought that any neuropsychological differences were associated with clinical and demographic variables. Nineteen studies met the inclusion criteria. Most findings pertained to AN diagnostic subtypes, for which there was scant robust evidence of any neuropsychological differences nor clinical or demographic associations. Using robust statistical techniques, profiles of neuropsychological performance have been discovered, though none that related to clinical or demographic variables. Neuropsychological research into AN diagnostic subtypes is hampered by poor sample sizes and inappropriate statistical techniques. Further research into AN subtype neuropsychological differences appears unwarranted. AN patients can be grouped into neuropsychological clusters, but at present these clusters seem theoretically interesting rather than clinically useful.

Introduction

Anorexia nervosa (AN) is an eating disorder that produces restricted eating, low body-mass index (BMI), fear of weight gain, and other unusual symptoms (Treasure, Claudino, & Zucker, 2010). Despite considerable research into treatment for anorexia nervosa (AN), outcomes are still poor, with relapse rates of up to 48% reported (Norring & Sohlberg, 1993; Papadopoulos, Ekbom, Brandt, & Ekselius, 2009). Pharmacological interventions are ineffective. For many years, the only effective psychological treatment was family-based therapy (Treasure et al., 2010), though recent guidelines have include eating-disorder-focused cognitive behavioural therapy, the Maudsley Anorexia Nervosa Treatment for Adults, and specialist supportive clinical management (National Institute for Health and Care Excellence, 2017). The struggle to find psychological treatments for AN has led researchers to consider neuropsychological approaches.

Anorexia and the Brain

Table 1.

AN Characteristics, Implicated Neural Structures, And Their Corresponding Functions

Symptom group	Eating disorder characteristic	Key neural structures	Hypothesised function
Core psychopathology	Morbid preoccupation with weight and shape	Somatosensory cortex	Body size evaluation
		Amygdala	Threat detection
		Frontal cortex	Information processing
	Distorted body image	Somatosensory cortex	Body size evaluation
		Hippocampus	Contextual memory
		Frontal cortex	Information processing

		Hypothalamus	Appetite and satiety
	Restricted food intake	Striatum	Reward value of eating
		Frontal cortex	Goal-directed behaviour
Associated features	Low self-esteem, shame and disgust	Insular cortex	Disgust
		Limbic system	Linking bodily experiences to feelings and thoughts
		Frontal cortex	Information processing
	Drive for thinness	Striatum	Reward
		Frontal cortex	Goal-directed behaviour
Common comorbid features	Obsessions and compulsions	Striatum	Compulsive behaviours
		Frontal cortex	Obsessional thoughts
	Anxiety	Amygdala	Threat detection
		Hippocampus	Contextual memory
	Depression	Hippocampus	Contextual memory
		Frontal cortex	Information processing
	Impaired visuo-spatial skills	Parietal cortex	Visual association processing
	Impaired executive functions	Frontal cortex	Cognitive inhibition
	Impaired empathy	Limbic system	Linking bodily experience to feelings and thoughts
		Frontal cortex	Information processing
	Anosognosia	Somatosensory cortex	Body evaluation
		Frontal cortex	Information processing
	Raised pain threshold	Thalamus	Pain processing
		Insular cortex	Pain processing
		Somatosensory cortex	Body evaluation

Note. AN = anorexia nervosa. Adapted from (Nunn, Frampton, Gordon, & Lask, 2008).

There are numerous characteristics of AN, associated neuroanatomical structures and neuropsychological functions (Table 1). The complex nature of AN is conveyed by its disparate features and their association with regions throughout the brain, from neocortex to diencephalon. Evidence of neuropsychological dysfunction in AN is considered below.

Intelligence. A meta-analysis of 30 studies involving 849 AN participants found AN patients had IQs 5.9 points greater than controls (Lopez, Stahl, & Tchanturia, 2010). Premorbid intelligence, as measured by the National Adult Reading Test (Nelson, 1982), correlated ($r = 0.56$, $p = 0.07$) with body mass index (BMI), raising the prospect that neuropsychological variables may associate with important clinical outcomes, though the correlation was performed using only a one-tailed test. An equivalent correlation was not significant in IQ measured with the Wechsler intelligence tests, which tap into fluid intelligence. Executive functions are thought to overlap with fluid intelligence and have been the subject of substantial research in AN.

Central coherence. The balance between processing the detail vs. gestalt of incoming stimuli is known as central coherence (Happé & Frith, 2006). Weak central coherence prioritises processing stimuli detail over the whole. Patients with AN are thought to possess weak central coherence (Lena, Fiocco, & Leyenaar, 2004; Carolina Lopez et al., 2008), which may manifest as an obsessional focus on weight and caloric intake at the expense of the bigger picture of their nutritional intake and long-term physical health.

Compared with healthy controls, weight-recovered AN patients showed improved performance on local processing and impaired performance on global processing (Lopez, Tchanturia, Stahl, & Treasure, 2009). That impairment endured after weight restoration suggests it may pre-date the onset of

starvation behaviour. Supporting evidence of the weak central coherence theory has also been found using the Fragmented Pictures Task (Harrison, Tchanturia, & Treasure, 2011), Rey Copy and Embedded Figures Test (Lopez et al., 2008), and Matching Familiar Figures Test (Southgate, Tchanturia, & Treasure, 2005).

Executive functioning. Cognitive functions that facilitate goal-directed activity are known as executive functions. They comprise working memory; inhibitory control; cognitive flexibility; and higher-order functions such as reasoning, problem-solving, and planning (Diamond, 2013). Patients with AN display impaired inhibition of obsessive thoughts about body shape and weight (Shafran & Somers, 1998). Adjusting cognitive activity consequent to environmental or goal changes requires set-shifting, impairments of which have been reported in patients with AN, recovered patients, and healthy sisters of siblings with AN (Roberts, Tchanturia, Stahl, Southgate, & Treasure, 2007; Tchanturia, Campbell, Morris, & Treasure, 2005). Some have theorised that set-shifting deficits underlie the compulsivity, rigidity, and perfectionism seen in AN patients (Tchanturia et al., 2004; Tchanturia, Morris, Surguladze, & Treasure, 2002).

Visuospatial processing and visual memory. The ability to perceive location, size, and features of visual stimuli requires visuospatial processing (Rauch & Savage, 1997). Compared to healthy controls, AN patients reliably show visuospatial weakness that persists after refeeding (Lena et al., 2004). Patients with AN also perform worse than healthy controls on the Rey Complex Figure Test (Key, O'Brien, Gordon, Christie, & Lask, 2006), though as the Rey test is also used to measure central coherence it may be difficult to discern which of the two functions contributed most to the poorer performance.

Patients with AN have visuospatial memory impairments and temporal lobe hypofusion (Key et al., 2006; Lask et al., 2005) and it has been suggested these deficits may be linked to the body-image distortion often present in AN (Favaro et al., 2012). Spatial recognition memory may also be impaired (Fowler et al., 2006). Patients with AN performed poorer than healthy controls on the Rey Copy and Rey Recall tests of immediate and delayed visual recall (Sherman et al., 2006). Visual memory impairments have not been found to improve at treatment follow up (Mikos et al., 2008), suggesting memory impairments may pre-exist self-starvation behaviour.

AN Subtypes

Diagnostically, there exist two AN subtypes: restricting (AN-R) and binge/purging (AN-BP). AN-R and AN-BP patients use different strategies for maintaining low weight: one avoids food, the other eats but purges, e.g., via vomiting or laxatives (American Psychiatric Association, 2013). Research suggests that AN-BP patients are more impulsive, affectively labile, and sensitive to interpersonal conflict than AN-R patients (Casper, Eckert, Halmi, Goldberg, & Davis, 1980; Garfinkel, Moldofsky, & Garner, 1980; Strober, 1981). Compared to AN-R patients, AN-BP patients are also thought to experience greater anxiety, depression, extroversion, and sexual activity; engage in substance misuse, theft, and self-harm; attempt suicide; respond less well to treatment; be more likely receive a personality disorder diagnosis; experience familial conflict; and report family psychiatric problems (Peat, Mitchell, Hoek, & Wonderlich, 2009; Pryor, Wiederman, & McGilley, 1996). AN-BP patient prognosis is worse than AN-R (Casper et al., 1980).

Diagnostic fluidity is a significant problem in the field of eating disorders. After 30 months, only 33% of 192 women diagnosed with AN, BN, or eating

disorder (ED) not otherwise specified (EDNOS) retained their original diagnosis (Milos, Spindler, Schnyder, & Fairburn, 2005). Symptom remission was a minor factor in this difference, with only 20% of the sample symptom-free at follow-up. The same fluidity is apparent between the AN-R and AN-subtypes, with a systematic review reporting rates of crossover from AN-R to AN-BP of 30-64% and from AN-BP to AN-R of 18-44% over 7-15 years (Peat et al., 2009). Whether the putative clinical differences in AN-R vs AN-BP are due to neuropsychological differences has received little attention, with much of the neuropsychological research in this area grouping the two subtypes – and even AN, BN, and EDNOS – together under the single category of AN or ED (e.g. Billingsley-Marshall et al., 2013; Israel et al., 2015). Given the neuropsychological features of AN, perhaps neuropsychologically-based classifications would be more valid and lead to more effective treatments.

Neuropsychological Profiles

Researchers have started to consider whether there may be more than one neuropsychological profile of AN. Encouraged by success in the psychosis field (Gilbert et al., 2014), AN researchers have sought to identify replicable neuropsychological profiles that can be correlated to clinical and demographic variables such as age, BMI, or treatment outcomes (Rose, Frampton, & Lask, 2012). Cluster analyses of AN neuropsychological performance have also revealed clusters (Renwick et al., 2015; Rose et al., 2016), though not ones that relate to clinical or demographic variables.

Not all researchers have found evidence for neuropsychological profiles in AN. Bentz et al. (2017) compared set-shifting, central coherence, processing speed, working memory, sustained attention, verbal memory, and verbal abstraction between first-episode AN patients, recovered AN patients, and

healthy controls. They found no difference in any of the neurocognitive abilities between the three groups. Bentz et al. called for a systematic review of the neuropsychological profile literature and greater standardisation of neuropsychological assessment between studies.

If neurocognitive profiles of AN or AN subtypes can reliably be identified and associated with clinical factors such as BMI, cognitions, treatment adherence and outcomes, we can improve our understanding of the disorder. We could develop neuropsychological profile-informed, personalised treatment pathways. Clinically-relevant neurocognitive profiles could form the basis of brain-directed treatments, e.g. cognitive-remediation therapy (Tchanturia, Davies, & Campbell, 2007). For example, the starvation strategy of constant restraint that characterises AN-R is highly cognitively demanding. AN-R patients may have greater neuropsychological capacity for inhibiting thoughts and behaviours that tempt them to eat than AN-BP patients. Conversely, AN-BP patients may have poorer inhibitory capacity than AN-R patients, which makes them unable to sustain the cognitive effort required to constantly stare themselves. The result is vacillation between food avoidance when inhibition capacity is adequate and food binges when inhibition is overwhelmed. The above literature suggests neuropsychological profiles may be present, but it is unclear whether they relate to clinically useful variables that might inform treatment. To discover if they do, a systematic review of the literature was conducted.

Research Questions

The following research questions (RQs) will be answered:

RQ1) Do patients with AN-R and AN-BP display different neuropsychological profiles?

RQ2) Are there observable clusters of neuropsychological functioning in patients with AN?

RQ3) Do differences observed in 1) and 2) correlate to any clinically relevant variables?

Method

A systematic review was conducted to answer the research questions.

Search Strategy

The review used the following databases: Ovid Medline, Embase, PsycINFO, Web of Knowledge, and Scopus. The references of all included papers were manually searched. Grey literature (i.e. thesis dissertations, technical reports, conference proceedings and authors) were searched using ProQuest Dissertations & Theses Global and the above databases.

Search Terms

Truncations will be used to capture the broadest usages of search terms. Separate Medical Subject Heading terms will be used with the Ovid databases (Ovid Medline, Embase and PsycINFO). Rather than searching exhaustively for every neuropsychological function and their synonyms, which could number in the hundreds, an umbrella approach was used. Neuropsychological functions were searched using truncations such as neuropsychology* and neurocognit*. A proximity operator (e.g. NEAR/2 in the example below) was used to link neuropsychological terms to the concept of impairment or difference (Table 2).

Table 2.

Literature Search String

#	Domain	Operator	Terms
1	Anorexia nervosa	AND	Anorex*
2	Endophenotype	OR	Endophenotype*
3	Executive functions	OR	Executive funct*
4	Neuropsychological functions	OR	Neuropsycholog*, neurocognit*, neuroscien*, cognitive
5	Performance differences	NEAR/2 or ADJ/2	abilit* OR performance OR function* OR reduc* OR discriminat* OR impair* OR dysfunc* OR deficit* OR difficult* OR weak* OR profile* OR cluster* OR subtype*

Note. Domain 5 was paired with each term from domain 4. Domains 5, 4, 3, and 2 were combined using the OR operator and paired with the anorexia domain 1 with the AND operator.

Screening Criteria

Papers were screened at two passes. First pass reviewed paper titles and abstracts against inclusion criteria outlined in Table 3. The second pass applied the same criteria to the full text of papers. A subsample of six papers was peer-reviewed to assess interrater reliability, which was high (r = 1.00).

Table 3.

Literature Review Inclusion Criteria

Criteria	Inclusion	Exclusion
Participants	Participants with anorexia nervosa diagnosed with the Diagnostic and Statistical Manual or Great Ormond Street Hospital guidelines (Bryant-Waugh, 2000)	AN mixed with ED (BN or EDNOS) such that direct comparison between AN without confounds is impossible.
Exposure	Comorbidities permissible but not required	Past or recovered AN only.
Comparison	Compare subtypes of AN (restricting vs. binge-purge) OR Compare clusters of neuropsychological functioning (e.g. AN group with strong	

	memory vs AN group with weak memory)	
Outcome	Test of neuropsychological functioning (e.g. memory, executive function, attention, visual processing)	Neuropsychological performance derived from questionnaire or self-report only.
Study Design	Cross sectional OR longitudinal, between-subjects comparison.	Qualitative

Evaluative Criteria

The quality of included papers was evaluated with a modified Newcastle-Ottawa Scale (NOS; Wells et al., 2001). The NOS adapted for cross-sectional studies (NOS-CS) assesses three characteristics: 1) selection, 2) comparability, and 3) outcome (Modesti et al., 2016). Papers are scored (maximum = 10) for providing information about sample size, non-respondents, between-subject control, and statistical quality (Appendix A). A subsample of three papers was peer-reviewed to assess interrater reliability, which was high (r = 0.94).

Results

Nineteen papers passed all eligibility criteria and were included in the review, down from an initial 1,243 database results (Figure 1). For brevity, results are described using study numbers (Table 4). Effect sizes (Cohen's d) were calculated for papers that did not report them (n = 12). A glossary is included to aid readability (Appendix B).

Participants

Study participants included adults (n = 15) and children (n = 9). Some studies included both (n = 5). Included papers reported results from 2,642 participants (M = 139.0, SD = 142.3). Excluding HCs and non-AN patients, there were 1,358 participants (M = 73.3, SD = 77.9). Specification of

participants was excellent, with all papers using either existing medical records, structured interview schedules, independent validation, or multidisciplinary decision making to confirm diagnoses.

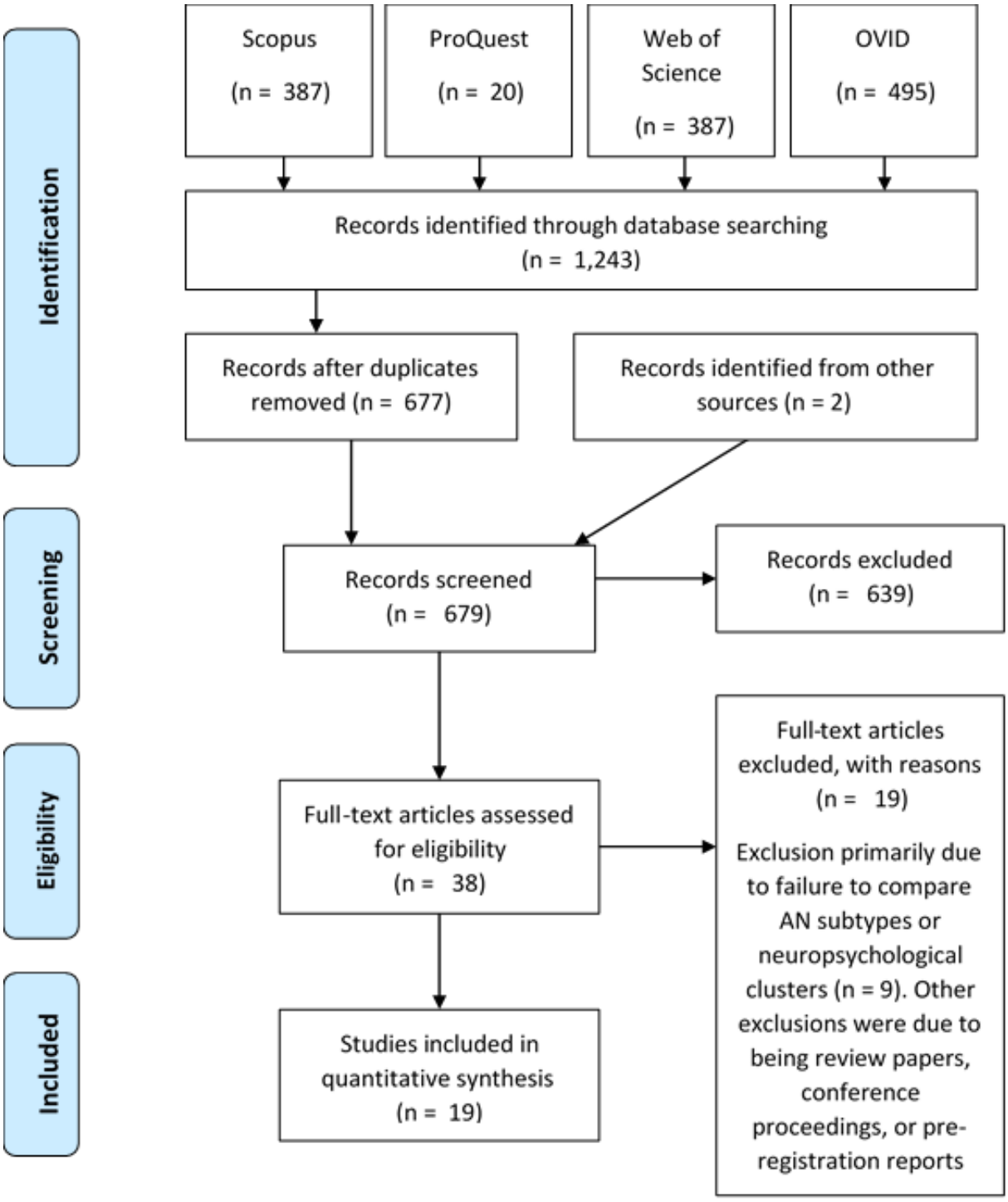


Figure 1. Study inclusion flow chart. ProQuest = ProQuest Dissertations and Theses. Adapted from Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group (2009).

Table 4.

Summary of Included Papers

#	Authors and date	*n*	Sample	Neuropsychological Function(s)	Measure(s)	Q	Findings
1	Abbate-Daga, Buzzichelli, Marzola, Amianto, & Fassino (2014)	153	78 AN-R, 16 AN-BP, 59 HC.	Set-shifting	WCST	7	**AN-subtype comparison and composite clustering.** NS AN-R and AN-BP re set-shifting. NS between impaired and intact set-shifting re age, education, and BMI. Impaired group greater depression (d = 0.94), interoceptive awareness (d = 0.82), maturity fears (d = 0.69), and impulse regulation (d = 0.74, ps < 0.001).
2	Andrés-Perpiña et al. (2011)	78	32 AN-R, 5 AN-BP, 41 HC.	IQ, immediate recall, delayed recall, visual perception, visual memory, auditory memory, processing speed, attention, cognitive flexibility, sorting, set-shifting, verbal fluency, interference control, central coherence	WISC-R, WMS-III, RCFT, RAVLT, TMT, WCST, COWAT, STP	7	**Composite clustering.** Impaired cognitive performers (scores on two or more tasks were > 2 SD less than the normative mean or scored 2 SD below predicted performance based on their intelligence) performed worse on visual memory and verbal memory. Significant difference between normal and impaired cognitive performance re BMI (d = 0.86) and trait anxiety (d = 0.90). All other demographic and clinical variables (age at onset, duration of illness, current age, depression, state anxiety, obsession severity, obsession interference, and attitude to eating) were equivalent (0.05 < d < 0.36).
3	Cavedini et al., (2004)	141	26 AN-R, 33	Decision making, set-shifting	IGT, WST, OAT, WCST	8	**AN-subtype comparison.** AN-R patients preferred disadvantageous deck selection (F = 11.57, P= 0.001). AN-BP patients showed no deck

			AN-BP, 82 HC				preference (p = 0.43). NS correlation between AN-R or AN-BP on IGT performance and BMI (0.02 < R^2 < 0.04), duration of illness in either and all other neuro tests.
4	Cavedini et al., (2006)	68	18 AN-R, 20 AN- BP, 30 HC	Decision making, set-shifting	IGT, WST, OAT	7	**AN-subtype comparison.** Significant time x IGT group performance re BMI for AN-BP (F = 11.12, p = 0.004) but not AN-R. NS between AN-R and AN-BP for IGT disadvantageous deck selection, WST, and OAT perseverative errors at admission and discharge. No effect sizes reported or calculable.
5	Dmitrzak-Weglarz et al., (2013)	105	46 AN-R, 14 AN-BP, 45 HC	Decision making, set-shifting	WCST	5	**AN-subtype comparison.** NS AN-R AN-BP re age, education, BMI (0.06 < ds < 0.11). NS AN-R AN-BP re WSCT perseverative errors, non-perseverative errors, correctly completed categories, conceptual level responses, set to first category (0.03 < ds < 0.14).
6	Galimberti, Martoni, Cavallini, Erzegovesi, & Bellodi (2012)	92	24 AN-R, 12 AN-BP, 16 BN, 40 HC	Motor inhibition, attentional set-shifting	CANTAB SST, IED	4	**Composite clustering.** NS between AN SST reaction time (RT) performers (bad, normal, good) re errors on go trials, mean correct RT on go trials, stop signal RT (motor inhibition); IED total errors (adjusted), interdimensional shift errors, and extradimensional shift errors or age, education, and duration of illness. Effect sizes not reported or calculable.
7	Giannunzio et al. (2018)	611	198 AN-R, 112 AN-BP, 301 HC	IQ, decision-making, set-shifting, abstraction, problem solving, perseveration, working memory,	BIT, WISC, WAIS, IGT, WCST, TMT, MWIT	8	**AN-subtype comparison and statistical clustering.** NS between adolescent (d = 0.25) or adult (d = 0.12) AN-R and AN-BP re IGT performance. Cluster analysis of IGT model parameters revealed a conservative and impulsive

				attention, motor speed, mental tracking			group, although the groups did not differ on any demographic variables (age or BMI) neuropsychological variables (IGT net, learning and risk; WCST perseveration and global; working memory; or TMT-B set-shifting scores) except the TMT-A measure of motor speed.
8	Godier et al., (2016) study 1	41	23 AN, 18 HC	Goal-directed behaviour, habit learning	SA	9	**AN-subtype comparison.** NS between AN-R and AN-BP re SA. Effect sizes not reported or calculable.
9	Herbrich, Kappel, Noort, & Winter (2018)	153	90 AN-R, 21 AN-BP, 63 HC	Fluid IQ, set-shifting, cognitive flexibility, central coherence	CFT-20-R, TMT, WCST, RCFT, GEFT	5	**AN-subtype comparison.** NS between AN-R and AN-BP re CFT-20-R (*d* = 0.24), TMT condition 4 (*d* = 0.09), WCST (*d* = 0.42), RCFT central coherence (*d* = 0.25), and GEFT (*d* = 0.38)
10	Nikendei et al. (2011)	99	34 AN-R, 19 AN-BP, 16 AN-WR, 30 HCs.	Logical short-term, long-term memory, working memory, selective attention, sustained attention, visual processing speed, lexical and semantic word fluency, cognitive flexibility, set-shifting	WMS-R, d2TA, RWFT, TMT	7	**AN-subtype comparison.** NS between AN-R and AN-BP re WMS-R (0.20 < *d*s < 0.54), d2TA (*d* = 0.10), RWFT (0.05 < *d*s < 0.15), and TMT-A (*d* = 0.26) and TMT-B (*d* = 0.58). AN-R slower (67.18s) on TMT-B than AN-BP (56.75s), though difference NS when corrected for 11 familywise comparisons.
11	Renwick et al. (2015)	100	44 AN-R, 33 AN-BP, 23 EDNOS-AN	Pre-morbid IQ, executive function, set-shifting, cognitive flexibility, visuo-spatial construction, central	NART, WCST, BSAT, RCFT, RMF	8	**Statistical clustering.** WCST global and perseverative errors, BSAT, RCFT copy, and RMF performance revealed a three-cluster solution. Cluster 1 displayed average to high performance, cluster 2 variable performance, and cluster 3 showed weak performance. Discriminant function

				coherence, emotional theory of mind			analysis confirmed that WCST global score and perseverative errors differentiated clusters 1 and 2 from 3, while BSAT was the most important for separating cluster 1 from 2. The clusters did not differ on any demographic or clinical variable: age, illness duration, BMI, eating disorder pathology, depression, anxiety, impairment, self-reported cognitive flexibility, obsessionality, beliefs about emotions, treatment adherence, relationship status, medication use, previous inpatient treatment, and treatment adherence ($0.01 < d$s < 0.70).
12	Roberts, Tchanturia, & Treasure (2016)	54	27 AN-R, 27 AN-BP	Set-shifting, central coherence	WCST, BSAT, TMT, HI, RCFT, GEFT.	5	**Composite clustering.** Compared to one/none impairments in set-shifting and/or central coherence (defined as 1 SD below mean), those with both were more severely ill ($d = 0.52$), prone to self-harm ($d = 0.58$), and ritualistic in their eating habits ($d = 0.48$).
13	Rose et al. (2016)	423	253 AN, 170 HC	IQ, memory, central coherence, cognitive flexibility, word productivity, inhibition, planning	WISC, WAIS, CFT-20-R RCFT; BSAT; DKEFS VF, TMT, CWI, and TT	7	**Statistical clustering.** Hierarchical cluster analysis revealed a three-cluster solution for AN patients and a two-cluster solution for HCs. AN Cluster 1 displayed low-average to average performance, AN cluster 2 produced weak visual and strong verbal performance and AN cluster 3 reported strong verbal performance. The HC clusters were equivalent to AN clusters 2 and 3. Discriminant function analysis revealed visual processing (RCFT) accounted for 62% of between-cluster variance and verbal fluency, verbal flexibility, and inhibition accounted for the remaining 38%. An equivalent HC analysis revealed a two-cluster

							solution that resembled AN clusters 2 and 3. NS between AN clusters on BMI, eating disorder pathology, depression, anxiety, and obsessionality ($0.01 < d < 0.31$).
14	Sato et al. (2013)	42	9 AN-R, 6 AN-BP, 27 HC	Set-shifting	WCST	6	**AN-subtype comparison.** NS between AN-R and AN-BP re WCST correct rate ($d = 0.65$), total error ($d = 1.42$), perseverative error ($d = 0.34$) and non-perseverative error ($d = 0.66$).
15	Tamiya et al. (2018)	109	22 AN-R, 18 AN-BP, 69 HC	IQ, processing speed, verbal fluency, attention, working memory, verbal learning, visual learning, reasoning/problem solving, social cognition	NART-J, MCCB-J	8	**AN-subtype comparison.** AN-BP significantly lower than AN-R on CPT-IT ($p = 0.01$). AN-R and AN-BP no correlation between neuropsychological performance and BMI (AN-R: $-0.20 \leq r \leq 0.30$, AN-BP: $-0.34 \leq r \leq 0.36$), minimum BMI (AN-R: $-0.36 \leq r \leq 0.08$, AN-BP: $-0.41 \leq r \leq 0.27$), or illness duration (AN-R: $-0.27 \leq r \leq 0.29$, AN-BP: $-0.11 \leq r \leq 0.51$).
16	Vall & Wade (2015)	177	18 AN-R, 9 AN-BP, 23 BN, 149 HC	Set-shifting, visual search, motor speed, perseveration	TMT	9	**AN-subtype comparison.** NS between AN-R and AN-BP ($0.129 < p\text{s} < 0.97$) re TMT Part A ($d = 0.33$), Part B ($d = 0.19$), Accuracy B ($d = 0.63$), and Part B–A ($d = 0.08$).
17	Van Autreve, De Baene, Baeken, van Heeringen, & Vervaet (2013)	77	31 AN-R, 20 AN-BP, 26 HC	Central coherence, set-shifting	WAIS-R BD, OA; WCST, TMT, TSP	6	**AN-subtype comparison.** AN-R significantly worse performance on OA than AN-BP ($M = 35.4$, $SD = 13.7$ vs. $M = 46.8$, $SD = 13.0$, $p = .006$), AN-R significantly worse than AN-BP on BD (M = 35.4, SD = 13.7 versus M 46.8, SD =13.0, $p < 0.001$). No effect sizes reported or calculable for mean comparisons. In AN-R, BD significantly correlated with education ($r = .40$, $p < .05$). In AN-BP, OA

							significantly correlated with trait anxiety (r = .55, p < .05).
18	Van Autreve et al. (2016)	44	16 AN-R, 13 AN-BP, 15 HC	Set-shifting	TSP	5	**AN-subtype comparison.** NS between AN-R and AN-BP re set-shifting repeat trials (d = 0.10), switch trials (d = 0.01), and switch costs (d = 0.19).
19	Yano et al. (2016)	75	9 AN-R, 17 AN-BP, 10 EDNOS, 39 HC	Inhibition	ASIT	5	**AN-subtype comparison.** Subgroups very small sample size. NS between AN-R and AB-BP re errors across congruent, incongruent, and neutral trial types (η^2 = 0.03) and RTs (η^2 = 0.04).

Note. # = study number, *n* = number of participants, D = design, Q = Quality (max = 10), AN = anorexia nervosa, AN-R = AN restricting subtype, AN-BP = AN bulimia subtype, HC = healthy control, EDNOS = eating disorder not otherwise specified, EDNOS-AN = EDNOS anorexia subtype, BN = bulimia nervosa, A = Adults, C = Children, CS = cross-sectional, L = longitudinal, NS = non-significant difference, ASIT = Arrow-space Interference Task, BACS-SC = Brief Assessment of Cognition in Schizophrenia – Symbol Coding test, BIT = Brief Intelligence Test, BD = Block Design, BSAT = Brixton Spatial Anticipation Test, BVMT-R = Brief Visuospatial Memory Test – Revised, CANTAB = Cambridge Neuropsychological Test Automated Battery, CFAN = Category Fluency – Animal Naming, CFT-20-R = Culture Fair Test 20-Revised, COWAT = Controlled Oral Word Association Test, CPT-IP = Continuous Performance Test – Identical Pairs, CWI = Colour-Word Interference, d2TA = d2 Test of Attention, DKEFS = Delis-Kaplan Executive Functional System, IED = Intra Dimensional/Extra Dimensional Shift Task, IGT = Iowa Gambling Task, GEFT = Group Embedded Figures Test, HI = Haptic Illusions, HVLT-R = Hopkins Verbal Learning Test – Revised, LNS = Letter-Number Span, MCCB-J = Japanese MATRICS Consensus Cognition

Battery, NAB-M = Neuropsychological Assessment Battery – Mazes, NART = National Adult Reading Test, NART-J = Japanese NART, MWIT = Memory With Interference Task, OA = Object Assembly, OAT = Object Alternation Test, RAVLT = Ray Auditory Verbal Learning Task, RCFT = Ray Complex Figure Test, RMF = Reading the Mind in Films Task, RWFT = Regensburg Word Fluency Test, SA = Slips-of-Action, SS = Spatial Span, SST = Stop Signal Task, STP = Stroop, TMT = Trail Making Task, TSP = a task switching paradigm, TT = Tower Tests, VF = Verbal Fluency, WAIS = Wechsler Adult Intelligence Scale, WCST = Wisconsin Card Sorting Task, WISC-R = Wechsler Intelligence Scale for Children Revised, WMS-R = Wechsler Memory Scale Revised, WMS-III = Wechsler Memory Scale III, WST = Weigl's Sorting Test. Organisational theme in bold.

Design

All but one study used a cross-sectional design. A single longitudinal paper reported differences between AN-R and AN-BP at admission and discharge from an inpatient treatment facility (#4). Those two cross-sectional comparisons were used for this review.

Quality

The quality of included studies was satisfactory (M = 6.5, SD = 1.57). Areas of weakness were failure to provide a sample size calculation (n = 18), unrepresentative cases or undetermined case representativeness (n = 10), and lack of control for important variables such as BMI, age, IQ, or education (n = 10).

Statistics

Sample sizes were generally poor and most papers underpowered. The average sample size of AN-R vs AN-BP comparisons was small (Mdn = 20.0). Assuming a medium effect size (Cohen, 1988), a total sample size of 158 would be required to detect a similar difference (α = 0.05, β = 0.80, d = 0.45). Only two of the nineteen papers (#7 and #13) were powered to detect such an effect.

Neuropsychological Tests

There was substantial heterogeneity among the neuropsychological tests used (n = 28). Given the number of tests used, only a brief description of each will be given (Table 5).

Table 5.

Included Neuropsychological Tests

Abrev.	Name	Author	Description	Function	S#
ASIT	Arrow-space Interference Task	Castel, Balota, Hutchison, Logan, & Yap (2007)	A modified Stroop task, in which the interference effect was produce by arrows pointing left or right, positioned on the left or the right of the screen	Inhibition	19
BACS-SC	Brief Assessment of Cognition in Schizophrenia –Symbol Coding test	Keefe et al. (2004)	From MCCB-J. A sheet of coded symbols must be deciphered into numerals using the legend provided	Attention and processing speed	15
BIT	Brief Intelligence Test	Colombo, Sartori, & Brivio (2002)	The Italian version of the National Adult Reading Test (NART)	Pre-morbid intelligence	7
BD	Block Design	Wechsler (1997)	From WAIS. Participants are required to use coloured cubes to replicate a complex geometric shape from a picture illustration	Visuo-spatial construction. Used in AN research as a measure of central coherence (C. Lopez, Tchanturia, Stahl, & Treasure, 2008).	17
BSAT	Brixton Spatial Anticipation Test	Burgess & Shallice (1997)	Participants observe the movement of a black dot among an array of white dots and predict its next location	Tests planning and inhibition.	11, 13
BVMT-R	Brief Visuospatial Memory Test–Revised	Benedict (1997)	A 2x3 matrix of abstract geometric designs is presented for 10 seconds. Participants then reproduce the designs and again after 25 minutes	Visual learning, immediate visual memory, delayed visual memory	15

CAN -TAB	Cambridge Neuropsychological Test Automated Battery	Robbins et al. (1994)	A neuropsychological test battery	Visual memory, visual attention, working memory, and planning	6
CFAN	Category Fluency – Animal Naming	Nuechterlein et al. (2008)	From MCCB-J. Equivalent to the category condition of the COWAT	Verbal production	15
CFT-20-R	Culture Fair Test 20-Revised	Weiß (2006)	A battery test of intelligence comprising eight subtests	Fluid intelligence	9, 13
COWAT	Controlled Oral Word Association Test	Benton, Hamsher & Sivan (1994)	Requires participants to produce words from a category or beginning with designated letter	Verbal production, though attention and working memory also contribute (Elias, Elias, D'Agostino, Silbershatz, & Wolf, 1997)	2
CPT-IP	Continuous Performance Test–Identical Pairs	Cornblatt, Risch, Faris, Friedman, & Erlenmeyer-Kimling (1988)	From MCCB-J. Participants respond by lifting their finding from a depressed key when two identical stimuli are presented consecutively	Sustained attention	15
CWI	Color-word Interference	Delis, Kaplan & Kramer (2001)	From DKEFS. A Stroop task in which participants read colour words printed in congruent or incongruent coloured inks. CWI contains an additional rule switching condition	Inhibition, set-shifting (flexibility)	13
d2TA	d2 Test of Attention	Brickenkamp (1972)	Participants mark target letters on a page while ignoring distractor letters	Visual scanning and sustained attention	10
DKEFS	Delis Kaplan Executive Functioning System	Delis, Kaplan, & Kramer (2001)	A battery of nine tests of executive functioning in the verbal and visual domain	Executive functioning	

IED	Intra Dimensional/Extra Dimensional Shift Task	Robbins et al. (1994)	From CANTAB. A computerised analogue of the Wisconsin Card Sorting Task (WCST). Extradimensional errors result from failing to learn the target stimulus after it switches from one dimension to another.	Learning and set-shifting	6
IGT	Iowa Gambling Task	Bechara, Damasio, Damasio, & Anderson (1994)	Participants select cards from four decks that either reward or penalise. Two decks are disadvantageous (i.e. high risk-reward) and reward fewer average credits	Set-shifting	3, 4, 7
GEFT	Group Embedded Figures Test	Witkin, Oltman, Raskin, & Karp (1971)	Requires participants to discern an embedded simple figure from within a complex figure over 18 trials	Central coherence	9, 12
HI	Haptic Illusions	Uznadze (1966)	Participants are habituated to two hand-held spheres of unequal size, then hold two spheres of equal size, which the hand previously habituated to the larger sphere nonetheless perceives as smaller than the sphere in other hand.	Set-shifting.	12
HVLT-R	Hopkins Verbal Learning Test–Revised	Benedict, Schretlen, Groninger, & Brandt (1998)	Participants to recall a list of 12 audially presented words in 3 trials, and again after a 20-25-minute interval (trial 4). Finally, participants are asked whether they recognise a 24-word list including the original 12 words.	Auditory short-term memory, auditory learning, delayed auditory recall, auditory recognition.	15
LNS	Letter-Number Sequencing	Wechsler (1991)	From WMS. A list of letters and numbers are presented. Participants must order the letters and numbers and repeat them aloud.	Auditory working memory	15

MCCB-J	Japanese Matrics Consensus Cognition Battery	Kaneda et al. (2013)	A Japanese version of the MCCB (Nuechterlein et al., 2008), an adult test battery. Comprises TMT, CFAN, CPT-IP, LNS, SS, HVLT-R, BVMT-R, and NAB-M	Processing speed attention/vigilance, working memory, verbal learning, visual learning, reasoning/problem-solving, social cognition	15
MWIT	Memory With Interference Task	Mondini, Mapelli, Vestri, Arcara, & Bisiacchi, (2011)	Participants are presented with three consonants to memorise while performing a simple addition to prevent rehearsal	Short-term memory	7
NAB-M	Neuropsychological Assessment Battery – Mazes	Stern & White (2003)	Participants must trace their way through a maze while visiting specific locations.	Executive functions, planning	15
NART	National Adult Reading Test	Nelson (1982)	Participants read words with irregular spellings	Pre-morbid intelligence	11
NART-J	Japanese NART	Matsuoka, Uno, Kasai, Koyama, & Kim (2006)	Participants read Japanese Kanji with irregular spellings	Pre-morbid intelligence	15
OA	Object Assembly	Wechsler (1997)	From WAIS. Participants compose an image from pieces, similar to a jigsaw puzzle.	Visuo-spatial and central coherence.	17
OAT	Object Alternation Test	Freedman (1990)	The OAT involves participants detecting the location of a penny from between two alternatives. The location changes after each correct response	Set-shifting and working memory	3, 4
RAVLT	Rey Auditory Verbal Learning Task	Rey (1964)	Participants must recall a list of audially presented words.	Short-term auditory memory	2
RCFT	Ray Complex Figure Test	Meyers & Meyers (1995)	The RCFT is a complex figure comprising 18 global and local components that participants must copy immediately and after a delay	Visuospatial processing, visual delayed recall, central coherence	2, 9, 11, 12, 13

RMF	Reading the Mind in Films Task	Golan, Baron-Cohen, Hill, & Golan (2006)	Using four-alternative forced-choice, participants identify what emotion is experienced by the subject of a film vignette	Theory of mind	11
RWFT	Regensburg Word Fluency Test	Aschenbrenner, Tucha, & Lange, (2000)	A word production test equivalent to the COWAT	Verbal production	10
SA	Slips-of-Action	Godier et al. (2016)	Participants learn to discriminate between high- and low-value stimuli under differing conditions that require them to inhibit previously learnt responses	Habit learning, set-shifting, and inhibition	8
SS	Spatial Span	Wechsler (1991)	From WMS. The examiner points sequentially to some of 10 patterned blocks. Participants must reproduce the same pattern from memory.	Visual working memory	15
SST	Stop Signal Task	Logan, Cowan, & Davis (1984)	The SST requires participants to press a button in response to a visual stimulus and to withhold the same response upon hearing an auditory cue present in 50% of trials	Motoric inhibition	6
STP	Stroop	Stroop (1935)	Equivalent to the CWI without the final rule-switching condition.	Selective attention, cognitive flexibility inhibition, interference	2
TMT	Trail Making Task	Reitan (1958)	The TMT tasks participants to draw a line between stimuli on a page. In TMT-B, the stimuli are from two sets (letters and numbers) and must be connected in ascending order while alternating between the two sets. TMT-A is a motor	Set-shifting, sustained attention, and alternate attention	2, 7, 9, 10, 12, 13,

			speed control condition where all the stimuli to be connected are from the same set (i.e. letters).		16, 17
TSP	Task switching paradigm	Jersild (1927)	Participants either categorise a presented number as odd vs. even or as large vs. small	Set-shifting and cognitive flexibility	17, 18
TT	Tower Test	Shallice, Broadbent, & Weiskrantz (1982)	In the TT, participants must move rings with different diameter between three poles to build a conical tower while not placing any larger ring over a smaller one	Planning, rule learning, inhibition	13
VF	Verbal Fluency	Delis, Kaplan, & Kramer (2001)	Word production task equivalent to COWAB.	Verbal production, though attention and working memory also contribute (Elias et al., 1997)	13
WAIS-R & WAIS-III	Wechsler Adult Intelligence Scale	Wechsler (1981) and Wechsler (1997)	Adult test batteries	General intelligence	7, 13, 17
WCST	Wisconsin Card Sorting Task	Grant & Berg (1948)	Sort cards into four decks according to an unknown rule that switches at set intervals	Set-shifting, planning, inhibition	1, 2, 3, 5, 7, 9, 11, 12, 14, 17
WISC-R	Wechsler Intelligence Scale for Children Revised	Wechsler (1991)	A child and adolescent test battery	General intelligence	2

WMS-R & WMS-III	Wechsler Memory Scale	Wechsler (1987) and Wechsler (1991)	Adult test batteries	Memory	2, 10
WST	Weigl's Sorting Test	Weigl (1941)	A sorting task equivalent to the WCST	Set-shifting, planning, inhibition	3, 4

Note. Abbrev. = abbreviation, S# = Study number. Function determined by reference to author's description or a neuropsychological test compendium (Strauss, Sherman & Spreen, 2006)

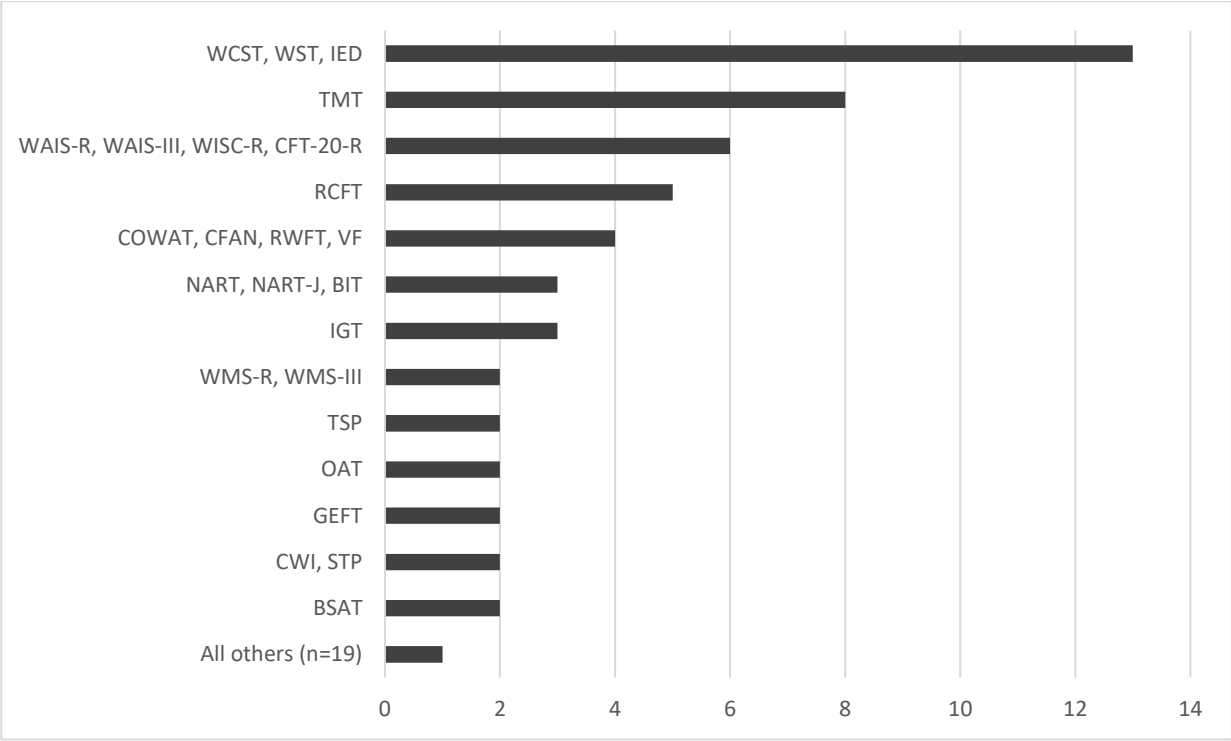


Figure 2. Frequency of neuropsychological test use. BIT = Brief Intelligence Test, BSAT = Brixton Spatial Anticipation Test, CFAN = Category Fluency – Animal Naming, CFT-20-R = Culture Fair Test 20-Revised, COWAT = Controlled Oral Word Association Test, CWI = Colour-Word Interference, IED = Intra Dimensional/Extra Dimensional Shift Task, IGT = Iowa Gambling Task, GEFT = Group Embedded Figures Test, NART = National Adult Reading Test, NART-J = Japanese NART, OAT = Object Alternation Test, RCFT = Ray Complex Figure Test, RWFT = Regensburg Word Fluency Test, STP = Stroop, TMT = Trail Making Task, VF = Verbal Fluency, WAIS = Wechsler Adult Intelligence Scale, WCST = Wisconsin Card Sorting Task, WISC = Wechsler Intelligence Scale for Children, WMS = Wechsler Memory Scale, WST = Weigl's Sorting Test.

The WCST-style tests, TMT, intelligence batteries, RCFT, and COWAT-style tests were the most popular, measuring set-shifting, inhibition, intelligence, central coherence, and verbal production respectively (Figure 2). Many tests appeared only once, though most of them examined the same neuropsychological functions as the more commonly used tests.

Critical Appraisal

The included studies fell into three types: comparisons between AN-R and AN-BP (RQ1), composite neuropsychological clustering (RQ2), and statistical neuropsychological clustering (RQ2). Composite clustering and statistical clustering are described separately because the two methodological approaches are not equally robust.

AN-subtype Comparison

These studies (*n* = 14) compared the performance of patients with AN-R and AN-BP. They contribute to answering RQ1 – that there are neuropsychological differences between AN-R and AN-BP – and RQ3 – that those differences are clinically meaningful.

Set-shifting. Many papers (*n* = 11) examined set-shifting between AN-R and AN-BP (#1, #3, #4, #5, #7, #9, #10, #14, #16, #17, #18), which was typically measured using the IGT, WST, OAT, TMT, or WCST (Table 5). Results generally provide weak or no evidence for differences between the two groups. Support for RQ1 was found in a single paper that reported AN-R patients preferring disadvantageous (i.e. high risk-reward) IGT deck selections, whereas AN-BP patients showed no preference (#3). While deck selections differed, overall performance did not, nor did performance correlate to the BMI or length of illness duration.

In support of RQ3, a tenuous association with a clinical outcome has been reported in one study. A significant interaction between IGT performance and BMI was reported (#4), with AN-BP patients in a high IGT-performance group displaying higher BMI post-treatment than low performers. Conversely, the AN-R group displayed no difference. The dichotomous IGT variable was created via a two-way split, which reduced the already low sample size (n = 38) and rendered the result high unreliable. Indeed, the supplied figures indicate the high IGT group gained more weight during treatment and ended treatment at a higher weight, irrespective of AN subtype. Additionally, substantial evidence (n = 9) suggests that, concerning set-shifting, RQ1 is unsupported. Reports of non-significant differences between AN-R and AN-BP include on the TMT (#9, #16), IGT (#4, #7), WST (#4), OAT (#4), TSP (#17), ASIT (#19), and WCST (#1, #3, #5, #9, #14). Importantly, study #7 was one of the studies with adequate power to detect a medium effect size difference; their null result severely undermines the evidence for RQ1.

Central coherence. A single paper found support for RQ1. Patients with AN-R performed more poorly than AN-BP on the WAIS-R BD and AO tests (#17). AN-R patient performance on BD was positively correlated with education (r = .40, p < 0.05) and AN-BP patient performance on OA was positively correlated with trait anxiety (r = .55, p < 0.05). However, the authors failed to correct for family-wise error from sixteen comparisons, meaning a 56% probability of type 1 error. Consequently, the two positive correlations should be treated with caution. No differences have been found between AN-R and AN-BP on the RCFT (#9) and GEFT (#9) measures of central coherence. As the sample size used by study #9 was superior (n = 111) to study #17 (n = 29), their null result further undermines support for RQ1 and RQ3.

Other functions. Five other papers were unsuccessful in supporting RQ1 (#7, #8, #9, 10, #15). Non-significant results have been observed on the MWI (#7) and WMS (#10) measures of memory; the NART, WISC, WAIS (#7), and CFR-20-R (#9) measures of intelligence; the SA (#8) measure of habit formation; the RWFT (#10) measure of verbal fluency; and the d2TA (#10) measure of selective and sustained attention (Table 5). The only positive finding was that AN-BP scored significantly lower than AN-R on the CPT-IT measure of attention, although performance did not correlate with any demographic or clinical variables (#15). This finding contradicts the null result of study #10. Given the modest sample size of both studies (#15, *n* = 40; #10, *n* = 40), both results may be unreliable.

In sum, papers testing RQ1 and RQ3 – that there are neuropsychological differences between AN-R and AN-BP and those differences are clinically meaningful – on set-shifting, central coherence, or other neuropsychological functions have failed to find any robust, convincing evidence. A plethora of negative results contrast the rare positive findings, which were universally underpowered and often statistically flawed.

Composite Clustering

Studies in this band (*n* = 4) split participants into subgroups based on their neuropsychological performance. Each of the four papers used different strategies for splitting neuropsychological performance, reducing their comparability. One study (#1) appears in both the composite clustering and AN-subtype organisational themes.

Set-shifting, motoric inhibition, and central coherence. When split into impaired (*n* = 30) and intact set-shifting ability (*n* = 64), as determined by

WCST score, AN-R and AN-BP patient distribution between the two groups was equivalent (#1), disconfirming RQ1. The impaired and intact set-shifting groups did not differ in terms of age, education, or BMI; they did differ on depression, interoceptive awareness, maturity fears, impulse regulation, and self-directedness ($0.69 < d < 0.94$), with the impaired set-shifting group scoring higher on each measure. That the intact set-shifting group was more depressed and differed on several cognitive dimensions provides some support for RQ2 and RQ3.

The other two studies that investigated set-shifting did so alongside other variables. In one study, using SST reaction times, patients were divided into bad, normal, and good SST performers (#6). Bad performers produced more extradimensional set-shifting errors than normal and good SST performers ($p = 0.039$) but were equivalent on other IED and SST metrics and clinical and demographic variables (Table 5). That the composite SST clusters differed on only one of six IED and SST metrics somewhat undermines the reliability and validity of the splitting procedure. The modest sample size ($n = 24$) in the AN cluster analysis also critically undermines the credibility of the results and their support for RQ2.

The final composite clustering paper used WCST, BSAT, TMT and HI to measure set-shifting alongside RCFT and GEFT measures of central coherence (#12). Patients were clustered according to impairment in either ($n = 41$) or both ($n = 11$) set-shifting and central coherence. When compared with demographic and clinical variables, patients with both impairments had lower lifetime BMI and were more severely ill, prone to self-harm, and ritualistic in their eating habits ($0.48 < d < 0.58$). The authors did not report *p*-values for any of the differences, so it is impossible to be confident that these observed differences were not due

to chance, which is especially problematic given the small and uneven sample sizes. Consequently, this study does not provide credible evidence to support either RQ2 or RQ3.

Impaired cognitive profiles. Patients have also been grouped according to impairments across their profile of neuropsychological test scores (#2). Patients with an impaired cognitive profile (n = 11) performed worse on the WMS and RCFT measures of visual memory and RAVLT measure of verbal memory compared to patients with a normal cognitive profile (n = 26), suggesting support for RQ2. No difference was observed between the two groups on the other 15 neuropsychological tests. Without family-wise correction, the uncontrolled type 1 error rate was substantial (54%). Patients with impaired cognitive profile had higher BMI (p = 0.023, d = 0.86) and trait anxiety (p = 0.028, d = 0.90) compared to patients with a normal cognitive profile. There was no clinical or demographic difference between the groups in any of the eight variables tested (Table 5). Due to the small, uneven sample sizes and failure to control the family-wise error rate, this study does not provide robust evidence to support either RQ2 or RQ3.

Notwithstanding the methodological limitations of data splitting, only a single study provided evidence that supported RQ2 and RQ3 (#1). The other studies either provided no confirmatory evidence or provided evidence so flawed that it does not support any research questions.

Statistical Clustering

The smallest methodological group (n = 3) contained papers that used robust statistical techniques to find clusters of neuropsychological test performance. One study from this group also reported AN-subtype comparisons

(#7); it is included in both the AN-subtype comparison and statistical clustering sections. Unlike the other two approaches used by the studies in this review, those that used statistical clustering techniques examined a range of neuropsychological functions. All scored well on the NOS-CS quality measure (M = 7.3 SD = 0.5), had acceptable or excellent sample sizes (M = 378, SD = 258.5), and used cluster analysis to determine the neuropsychological groups.

Using a hierarchical cluster analysis based on performance parameters from the IGT (#7), researchers discovered two symmetrical groups, one conservative, the other impulsive, providing support for RQ2. The conservative group displayed high choice constancy, low updating, and low motivation; they were less sensitive to losses, made prudent card selections, and were disinclined to update their selection strategy. The impulsive group was characterised by low choice constancy, high updating, and high motivation; their selections were influenced by recent feedback (not previous selections) more by gains than by losses and were susceptible to task fatigue. Analysis of the clusters revealed that the impulsive cluster was populated significantly more by adolescents than adults (50% vs. 22%, $p < 0.001$) in the healthy sample but not among AN patients (31% vs. 27%, $p = 0.493$), suggesting that whereas HC adolescent impulsivity maturates in adulthood, AN adolescent impulsivity may do so less. Among AN patients, after controlling for education, there was no difference between the conservative and impulsive cluster on most clinical or demographic variables or other neuropsychological performance (Table 5). The only significant difference was that the impulsive group completed the TMT-A significantly faster ($p = 0.008$) than the conservative group, demonstrating improved motor speed. The study failed to provide support for RQ3 but was

restricted to IGT-based clusters. Other researchers have clustered across a range of neuropsychological dimensions.

Hierarchical cluster analysis on WCST, BSAT, RCFT, and RMF revealed three clusters of performance (#11), supporting RQ2. Cluster 1 (*n* = 45) showed slight strength in set-shifting, mild weakness in central coherence and average performance on all other functions. Cluster 2 (*n* = 38) showed strong set-shifting and emotional theory of mind, and weak central coherence and cognitive flexibility. Cluster 3 (*n* = 17) demonstrated weak set-shifting, central coherence, and emotional theory of mind, with average performance on cognitive flexibility. The three clusters did not differ on any demographic or clinical measures or IQ (Table 5), which weakens support for RQ3. The distribution of eating disorder subtypes was equivalent between the three clusters, undermining RQ1.

The final paper in the statistical clustering methodological group (#13) used k-means cluster analysis to group patients and HCs across 15 neuropsychological variables covering the full profile of functions considered affected by AN: IQ, memory, central coherence, set-shifting, word production, inhibition, and planning. A three-cluster solution best discriminated AN patient neuropsychological performance. Cluster 1 (*n* = 49), contained patients with weak visual memory and inhibition and average central coherence, verbal fluency, verbal flexibility, and planning. Cluster 2 (*n* = 83), demonstrated weak visual memory, strong verbal fluency and verbal flexibility, and otherwise average performance. Cluster 3 (*n* = 121), performed well on verbal fluency and verbal flexibility and average in all other tests. The three clusters differed on all neuropsychological tests with extremely large effect sizes ($0.15 < \eta_p^2 < 0.419$), except for the TT ($\eta_p^2 = 0.014$), BSAT ($\eta_p^2 = 0.053$), and RCFT recognition ($\eta_p^2 = 0.051$), which had large, medium, and medium effect sizes respectively. The

identified clusters provide strong support for RQ2. Unfortunately, none of the three AN clusters differed on any clinical or demographic variable (Table 5), which provides strong evidence that RQ3 is not supported.

Discussion

Where differences between AN subtypes or neuropsychological profiles have been found, the implications for RQ3, that differences are associated with clinical or demographic variables, will be considered.

Neuropsychological Differences Between AN Subtypes

Overall, there was little evidence to support RQ1. Differences between AN-R and AN-BP performance were observed in one paper, but they did not relate to any clinically meaningful variables, such as BMI or illness duration (#3), so fail to provide support for RQ3. The only well-powered paper addressing RQ1 found impulsive (as defined by IGT performance) AN patients possessed greater motoric speed than conservative patients (#7). No support was found for RQ3. Nor was RQ1 supported by any of the ten other studies that investigated set-shifting in AN subtypes.

Only weak or null evidence was found to support RQ1 regarding central coherence difference between AN-R and AN-BP. One paper reported AN-R performing poorly compared with AN-BP on tests of central coherence (#17). Although this difference correlated with clinically relevant variables (education and trait anxiety), statistical flaws severely undermine reliability and support for RQ3. Co-occurring set-shifting and central coherence impairments may correlate with lower lifetime BMI and illness severity (#12) but again, support for RQ3 is hindered by statistical flaws.

Papers investigating memory, intelligence, and verbal fluency have all failed to produce evidence to support RQ1 difference between AN-R and AN-BP. Contradictory evidence has been produced by two papers (#10, #15). One found a null result, the other found that AN-BP patients have worse attention than AN-R patients (#15), though the study was underpowered and failed to correlate AN subtype neuropsychological differences with clinical variables, thereby not supporting RQ3.

Neuropsychological Profiles Within AN

One paper found weak support for RQ2 and RQ3 (#4) but the comparison was a severely underpowered ANOVA that utilised mean splitting, a technique that is indefensible in such research. The findings of set-shifting studies within the composite clustering theme were similarly unimpressive, each failing to support either RQ2 or RQ3 (#1, #6, #12). Poor performance across several neuropsychological measures has been found to be associated with visual and verbal memory deficits (#2), though statistical and sample size shortcomings undermine its support for RQ2. The study also failed to support RQ3.

Using statistical clustering methods, robust support has been found for RQ2 differences in neuropsychological function within AN. In one study, set-shifting, central coherence, emotional reasoning performance produced three clusters, although the clusters did not differ from each other on clinical and demographic measures nor IQ (#11). AN subtypes were also equivalent between the three clusters, contrary to RQ1. Another study performed a similar analysis using 423 participants (253 patients); however, it produced similarly disappointing results (#13). While RQ2 was supported by identifying neuropsychological clusters, those clusters did not correlate with any clinical or

demographic variables. The failure to link observed clusters with clinical and demographic variables, given the excellent sample size and robust statistical methodology, suggest that RQ3 is not supported.

In sum, there is insufficient robust and credible evidence among any of the studies to support RQ1. The evidence suggests there are no reliable neuropsychological differences between the AN-R and AN-BP subtypes. Composite clustering studies, due to methodological and statistical shortcomings, fail to provide compelling evidence for RQ2. Statistical clustering methods do provide credible evidence that neuropsychological differences exist within AN (RQ2); however, they fail to evidence RQ3 that such differences relate to clinical or demographic variables. While interesting, detecting neuropsychological differences between AN patients is insufficient if researchers cannot relate them to other variables. Enormous neuropsychological heterogeneity is observable within the general population (Wechsler, 2010). Any differences observed in AN patients must be accompanied by associated differences in clinical or demographic outcomes, otherwise researchers are simply measuring the neuropsychological heterogeneity present in all humans and wrongly extrapolating significance that is not warranted.

Review Results in Context

A significant challenge for researchers attempting to use neuropsychological tests to understand anorexia is their validity. Neuropsychological test authors claim that their test interrogates one or other function. They correlate their test with other measures thought to measure the same or similar functions – tests themselves validated in the same circular way. Once published, others use the test and may report that it also measures

additional functions. For example, the authors describe the BSAT as a measure of planning and inhibition (Burgess & Shallice, 1997). It has been used by AN researchers to assess set-shifting (Renwick et al., 2015) and central coherence (Rose, Davis, Frampton, & Lask, 2011). Test performance may also be influenced by sensory domain, in the case of the BSAT, vision. Assumptions are made about the equivalence of findings obtained in using a test in one sensory domain to other sensory domains, such as audition. Some tests (e.g. WCST) provide multiple performance metrics, which are used inconsistently by different research groups. These variables permit substantial researcher degrees of freedom and, consequently, it is extremely challenging to interpret the results of such heterogeneous methodology. This point was made by Rose et al. (2011) when they wrote to call for a standardised neuropsychological assessment battery of AN patients. In writing this report, I began tabulating the frequency of neuropsychological functions tested by the included papers, but the volume and inconsistency made it uninterpretable. The debate regarding what functions are really tested in each neuropsychological test is beyond the scope of this review (see Lezak & Howieson, Bigler & Tranel, 2012).

Methodological flaws were present in many of the included papers. All but two were severely underpowered to detect differences between AN-subtypes, though in most papers subtype comparisons were a supplementary research question. Sample sizes were frequently uneven, with AN-R much more commonly included than AN-BP, reflecting the unequal ratio between these diagnostic subtypes in patients. The entire organisational theme of composite clustering can be critiqued, as the method of artificially creating dichotomous or ordinal groups from continuous variables has been severely criticised. Any split is sample specific, meaning that, even with identical

methodology, two studies may produce low-high split groups that substantially overlap with one another purely due to sampling error (Pastor, Barron, Miller, & Davis, 2007). The method also spuriously inflates the similarity between participants within groups and the differences between groups, such that scores of one and ten in a ‘low’ group could be considered statistically more similar with each other than scores of eleven and twelve in a ‘high’ group, if those two groups are discriminated by the mean of ten. The statistical consequences of data splitting include information loss, reduced effect size and power, increased type 1 errors, inability to detect non-linear effects, and impeded cross-study comparison (MacCallum, Zhang, Preacher, & Rucker, 2002). MacCallum et al. conclude that splitting is almost never justified. The present review recommends its use within AN neuropsychological research should be abandoned.

Cluster analysis has been used to attempt to identify unobserved groups within neuropsychological test scores despite it being suboptimal for the task. Cluster analysis does not provide a statistic to quantify the fit of a model, leaving researchers to subjectively decide on model structure based on conventions and techniques such as eyeballing the data. It is also negatively affected by multicollinearity (Ketchen & Shook, 1996). Superior methods such as latent profile analysis (Muthén & Muthén, 1998-2012) should be used by researchers attempting similar analyses in the future. Latent profile analysis enables person-centred statistical clustering that provides a probability value for assignment to each latent group that is modelled. Additionally, the method also uses model fit statistics, thereby reducing the experimenter degrees of freedom and reducing bias (Berlin, Williams, & Parra, 2014).

Given all the null findings among the AN subtype and neuropsychological clusters literature, what sense are we to make of the robust findings that there is

neuropsychological dysfunction in AN patients (Harrison et al., 2011; Key et al., 2006; Lopez et al., 2010; Tchanturia et al., 2005)? One possible explanation is that our current measurements are inadequate to fully capture the phenomenology of AN. The demographic and clinical variables against which neuropsychological functioning has been compared include age, BMI (lifetime lowest, current, and pre- and post-treatment); illness duration; treatment adherence, medication use, education; trait and state depression, anxiety, obsessionality, and eating disorder pathology (eating restraint, weight concern, shape concern, and eating concern). It may be that researchers are not currently measuring the right variables or that the current variables are measured inappropriately. Measurement is an eternal problem for psychology: we are unable to directly measure much of the human experience and whatever we do measure is invariably contaminated with noise. The clinical and demographic variables that were measured do not currently include such things as history of adverse childhood experiences, attachment type, alexithymia, family functioning, or autistic features, each of which is associated with AN in some way (Bourke, Taylor, Parker, & Bagby, 1992; Gowers & North, 1999; Jaite et al., 2012; Ward et al., 2001; Zucker et al., 2007). The preponderance of null findings in this review suggests that AN subtypes may not help us to understand the process that produces pathological food avoidance. Perhaps the subtypes do nothing more than describe the two predominant strategies available to a patient with AN who wishes to starve themselves. Including the above variables may offer additional targets to correlate neuropsychological profiles with, but the cost of conducting such research is substantial and highly speculative given the current state of the literature. The neuropsychological clusters that have been found have not been correlated with demographic and clinical variables. They

may serve a function and perhaps may inform future work attempting to use neuropsychological clusters to identify genes important to the development of AN or that are activated or deactivated in an epigenetic process that may link to environmental effects such as family dysfunction or emotionally intolerant households.

Limitations

The principal limitation of this review was its breadth. The number of neuropsychological functions that are associated with AN is huge. The original search strategy adopted was to attempt to exhaustive list all the relevant functions, but this quickly became unmanageable. Attempting to capture all possible synonyms of a function and the countless ways a function might be specified – and in such a way that it would not be captured by a superordinate term – proved impossible. The alternative approach of capturing papers by using broad terms such as neuropsychological, neuroscience, neurocognitive successfully captured many papers without producing an overwhelming number of irrelevant results. Some papers may have been missed in using general search terms, though only two additional papers were found among the references of included papers, which indicates that this was not a major problem.

Future Directions

The current literature suggests there is currently no compelling evidence of any neuropsychological differences between patients with AN-R and AN-BP. Any attempt to continue to explore the neuropsychological differences between the two subtypes should be abandoned. Attempts to identify clinically meaningful differences between AN patients based on neuropsychological

profiles have also been unfruitful. AN researchers should improve the sample sizes of their studies and cease methodological practices such as data splitting. Any future investigation of neuropsychologically-based clusters should be done using appropriate person-centred statistical methods such as latent profile analysis. If those attempts fail to yield clinically meaningful results, then they too should be abandoned, and research refocused from the brain onto the person.

Conclusion

This systematic literature review of putative hitherto undiscovered AN subtypes focused on current diagnostic subtypes (AN-R and AN-BP) and latent neuropsychological profiles. The findings from 19 papers failed to provide compelling evidence of any clinically meaningful difference between the AN subtypes nor between the neuropsychological profiles uncovered with cluster analysis. This failure suggests that, insofar as there are neuropsychological differences between patients with AN beyond those observable in the general population, these differences are not likely to enhance our ability to understand the disorder or find more effective treatments.

References

Abbate-Daga, G., Buzzichelli, S., Marzola, E., Amianto, F., & Fassino, S. (2014). Clinical investigation of set-shifting subtypes in anorexia nervosa. *Psychiatry Research*, *219*(3), 592–597. https://doi.org/10.1016/j.psychres.2014.06.024

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Washington, DC: Author.

Andrés-Perpiña, S., Lozano-Serra, E., Puig, O., Lera-Miguel, S., Lázaro, L., & Castro-Fornieles, J. (2011). Clinical and biological correlates of adolescent anorexia nervosa with impaired cognitive profile. *European Child and Adolescent Psychiatry*, *20*(11–12), 541–549. https://doi.org/10.1007/s00787-011-0216-y

Aschenbrenner, S., Tucha, O., & Lange, K. W. (2000). *The Regensburg Word Fluency Test.* Hogrefe: Göttingen.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Benedict, R. H. B. (1997). *Brief Visuospatial Memory Test – Revised.* Odessa, FL: Psychological Assessment Resources.

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test – Revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist*, *12*(1), 43–55. https://doi.org/10.1076/clin.12.1.43.1726

Benton, A. L., Hamsher, K., Rey, G. L., & Sivan, A. B. (1994). *Multilingual aphasia examination* (3rd ed.). Iowa City, IA: AJA Associates.

Bentz, M., Jepsen, J. R. M., Kjaersdam Telléus, G., Moslet, U., Pedersen, T., Bulik, C. M., & Plessen, K. J. (2017). Neurocognitive functions and social functioning in young females with recent-onset anorexia nervosa and recovered individuals. *Journal of Eating Disorders*, *5*, 5. https://doi.org/10.1186/s40337-017-0137-3

Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An Introduction to latent variable mixture modeling (Part 1): Overview and cross-sectional latent class and latent profile analyses. *Journal of Pediatric Psychology*, *39*(2), 174–187. https://doi.org/10.1093/jpepsy/jst084

Billingsley-Marshall, R. L., Basso, M. R., Lund, B. C., Hernandez, E. R., Johnson, C. L., Drevets, W. C., … Yates, W. R. (2013). Executive function in eating disorders: the role of state anxiety. *The International Journal of Eating Disorders*, *46*(4), 316–321. https://doi.org/10.1002/eat.22086

Bourke, M. P., Taylor, G. J., Parker, J. D. A., & Bagby, R. M. (1992). Alexithymia in women with anorexia nervosa: a preliminary investigation. *The British Journal of Psychiatry*, *161*(2), 240–243. https://doi.org/10.1192/bjp.161.2.240

Brickenkamp, R. (1972). *Handanweisungen zum d2-Test [Manual for d2 Test of Attention].* Göttingen: Hogrefe Ltd.

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests.* Bury St Edmunds: Thames Valley Test Company.

Casper, R. C., Eckert, E. D., Halmi, K. A., Goldberg, S. C., & Davis, J. M. (1980). Bulimia. Its incidence and clinical importance in patients with anorexia nervosa. *Archives of General Psychiatry*, *37*(9), 1030–1035.

Castel, A. D., Balota, D. A., Hutchison, K. A., Logan, J. M., & Yap, M. J. (2007). Spatial attention and response control in healthy younger and older adults and individuals with Alzheimer's disease: Evidence for disproportionate selection impairments in the Simon task. *Neuropsychology*, *21*(2), 170–182. https://doi.org/10.1037/0894-4105.21.2.170

Cavedini, P., Bassi, T., Ubbiali, A., Casolari, A., Giordani, S., Zorzi, C., & Bellodi, L. (2004). Neuropsychological investigation of decision-making in anorexia nervosa. *Psychiatry Research*, *127*(3), 259–266. https://doi.org/10.1016/j.psychres.2004.03.012

Cavedini, P., Zorzi, C., Bassi, T., Gorini, A., Baraldi, C., Ubbiali, A., & Bellodi, L. (2006). Decision-making functioning as a predictor of treatment outcome in anorexia nervosa. *Psychiatry Research*, *145*(2), 179–187. https://doi.org/10.1016/j.psychres.2004.12.014

Colombo, L., Sartori, G., & Brivio, C. (2002). Estimation of intelligence quotient using the TIB (Intelligence Brief Test), *Giornale Italiano di Psicologia, 3*, 613–637. https://doi.org/10.1421/1256

Cornblatt, B. A., Risch, N. J., Faris, G., Friedman, D., & Erlenmeyer-Kimling, L. (1988). The continuous performance test, identical pairs version (CPT-IP): I. new findings about sustained attention in normal families. *Psychiatry Research*, *26*(2), 223–238. https://doi.org/10.1016/0165-1781(88)90076-5

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System.* London: Pearson Assessment.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dmitrzak-Weglarz, M., Skibinska, M., Slopien, A., Tyszkiewicz, M., Pawlak, J., Maciukiewicz, M., … Hauser, J. (2013). Serum neurotrophin concentrations in polish adolescent girls with anorexia nervosa. *Neuropsychobiology*, *67*(1), 25–32. https://doi.org/10.1159/000343500

Elias, M. F., Elias, P. K., D'Agostino, R. B., Silbershatz, H., & Wolf, P. A. (1997). Role of age, education, and gender on cognitive performance in the Framingham Heart Study: community-based norms. *Experimental Aging Research*, *23*(3), 201–235. https://doi.org/10.1080/03610739708254281

Favaro, A., Santonastaso, P., Manara, R., Bosello, R., Bommarito, G., Tenconi, E., & Di Salle, F. (2012). Disruption of visuospatial and somatosensory functional connectivity in anorexia nervosa. *Biological Psychiatry*, *72*(10), 864–870. https://doi.org/10.1016/j.biopsych.2012.04.025

Fowler, L., Blackwell, A., Jaffa, A., Palmer, R., Robbins, T. W., Sahakian, B. J., & Dowson, J. H. (2006). Profile of neurocognitive impairments associated with female in-patients with anorexia nervosa. *Psychological Medicine*, *36*(4), 517–527. https://doi.org/10.1017/S0033291705006379

Freedman, M. (1990). Object alternation and orbitofrontal system dysfunction in Alzheimer's and Parkinson's disease. *Brain and Cognition*, *14*(2), 134–143. https://doi.org/10.1016/0278-2626(90)90025-J

Galimberti, E., Martoni, R. M., Cavallini, M. C., Erzegovesi, S., & Bellodi, L. (2012). Motor inhibition and cognitive flexibility in eating disorder subtypes. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *36*(2), 307–312. https://doi.org/10.1016/j.pnpbp.2011.10.017

Garfinkel, P. E., Moldofsky, H., & Garner, D. M. (1980). The heterogeneity of anorexia nervosa. Bulimia as a distinct subgroup. *Archives of General Psychiatry*, *37*(9), 1036–1040.

Garner, D. (1991). *Eating Disorder Inventory 2*. Psychological Assessment Resources: Odessa.

Giannunzio, V., Degortes, D., Tenconi, E., Collantoni, E., Solmi, M., Santonastaso, P., & Favaro, A. (2018). Decision-making impairment in anorexia nervosa: New insights into the role of age and decision-making style. *European Eating Disorders Review*, *26*(4), 302–314. https://doi.org/10.1002/erv.2595

Gilbert, E., Mérette, C., Jomphe, V., Emond, C., Rouleau, N., Bouchard, R. H., … Maziade, M. (2014). Cluster analysis of cognitive deficits may mark heterogeneity in schizophrenia in terms of outcome and response to treatment. *European Archives of Psychiatry and Clinical Neuroscience*, *264*(4), 333–343. https://doi.org/10.1007/s00406-013-0463-7

Godier, L. R., de Wit, S., Pinto, A., Steinglass, J. E., Greene, A. L., Scaife, J., … Park, R. J. (2016). An investigation of habit learning in anorexia nervosa. *Psychiatry Research*, *244*, 214–222. https://doi.org/10.1016/j.psychres.2016.07.051

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The ‘reading the mind in films’ task: Complex emotion recognition in adults with and

without autism spectrum conditions. *Social Neuroscience*, *1*(2), 111–123. https://doi.org/10.1080/17470910600980986

Gowers, S. G., & North, C. (1999). Difficulties in family functioning and adolescent anorexia nervosa. *The British Journal of Psychiatry*, *174*(1), 63–66. https://doi.org/10.1192/bjp.174.1.63

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*, 404-411.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Harrison, A., Tchanturia, K., & Treasure, J. (2011). Measuring state trait properties of detail processing and global integration ability in eating disorders. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry*, *12*(6), 462–472. https://doi.org/10.3109/15622975.2010.551666

Herbrich, L., Kappel, V., van Noort, B. M., & Winter, S. (2018). Differences in set-shifting and central coherence across anorexia nervosa subtypes in children and adolescents. *European Eating Disorders Review*, *26*(5), 499–507. https://doi.org/10.1002/erv.2605

Israel, M., Klein, M., Pruessner, J., Thaler, L., Spilka, M., Efanov, S., … Steiger, H. (2015). n-Back task performance and corresponding brain-activation patterns in women with restrictive and bulimic eating-disorder variants:

Preliminary findings. *Psychiatry Research: Neuroimaging*, *232*(1), 84–91. https://doi.org/10.1016/j.pscychresns.2015.01.022

Jaite, C., Schneider, N., Hilbert, A., Pfeiffer, E., Lehmkuhl, U., & Salbach-Andrae, H. (2012). Etiological role of childhood emotional trauma and neglect in adolescent anorexia nervosa: A cross-sectional questionnaire analysis. *Psychopathology*, *45*(1), 61–66. https://doi.org/10.1159/000328580

Jersild, A. T (1927). Mental set and shift. *Archives of Psychology, no. 89*. Retrieved from https://archive.org/details/mentalsetshift00jers

Kaneda, Y., Ohmori, T., Okahisa, Y., Sumiyoshi, T., Pu, S., Ueoka, Y., … Sora, I. (2013). Measurement and treatment research to improve cognition in schizophrenia consensus cognitive battery: Validation of the Japanese version. *Psychiatry and Clinical Neurosciences*, *67*(3), 182–188. https://doi.org/10.1111/pcn.12029

Keefe, R. S. E., Goldberg, T. E., Harvey, P. D., Gold, J. M., Poe, M. P., & Coughenour, L. (2004). The Brief Assessment of Cognition in Schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research*, *68*(2–3), 283–297. https://doi.org/10.1016/j.schres.2003.09.011

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal*, *17*(6), 441–458. Retrieved from JSTOR.

Key, A., O'Brien, A., Gordon, I., Christie, D., & Lask, B. (2006). Assessment of neurobiology in adults with anorexia nervosa. *European Eating Disorders Review*, *14*(5), 308–314. https://doi.org/10.1002/erv.696

Lask, B., Gordon, I., Christie, D., Frampton, I., Chowdhury, U., & Watkins, B. (2005). Functional neuroimaging in early-onset anorexia nervosa. *The International Journal of Eating Disorders*, *37 Suppl*, S49-51; discussion S87-89. https://doi.org/10.1002/eat.20117

Lena, S. M., Fiocco, A. J., & Leyenaar, J. K. (2004). The role of cognitive deficits in the development of eating disorders. *Neuropsychology Review*, *14*(2), 99–113. https://doi.org/10.1023/B:NERV.0000028081.40907.de

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment (5th ed.).* Oxford: Oxford University Press.

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(2), 276–291.

Lopez, C., Tchanturia, K., Stahl, D., & Treasure, J. (2008). Central coherence in eating disorders: a systematic review. *Psychological Medicine*, *38*(10), 1393–1404. https://doi.org/10.1017/S0033291708003486

Lopez, C., Stahl, D., & Tchanturia, K. (2010). Estimated intelligence quotient in anorexia nervosa: a systematic review and meta-analysis of the literature. *Annals of General Psychiatry*, *9*, 40. https://doi.org/10.1186/1744-859X-9-40

Lopez, C., Tchanturia, K., Stahl, D., Booth, R., Holliday, J., & Treasure, J. (2008). An examination of the concept of central coherence in women

with anorexia nervosa. *The International Journal of Eating Disorders*, *41*(2), 143–152. https://doi.org/10.1002/eat.20478

Lopez, C., Tchanturia, K., Stahl, D., & Treasure, J. (2009). Weak central coherence in eating disorders: A step towards looking for an endophenotype of eating disorders. *Journal of Clinical and Experimental Neuropsychology*, *31*(1), 117–125. https://doi.org/10.1080/13803390802036092

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40.

Matsuoka, K., Uno, M., Kasai, K., Koyama, K., & Kim, Y. (2006). Estimation of premorbid IQ in individuals with Alzheimer's disease using Japanese ideographic script (Kanji) compound words: Japanese version of National Adult Reading Test. *Psychiatry and Clinical Neurosciences*, *60*(3), 332–339. https://doi.org/10.1111/j.1440-1819.2006.01510.x

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and recognition trial: Professional manual*. Odessa: Psychological Assessment Resources.

Mikos, A. E., McDowell, B. D., Moser, D. J., Bayless, J. D., Bowers, W. A., Andersen, A. E., & Paulsen, J. S. (2008). Stability of neuropsychological performance in anorexia nervosa. *Annals of Clinical Psychiatry: Official Journal of the American Academy of Clinical Psychiatrists*, *20*(1), 9–13. https://doi.org/10.1080/10401230701844836

Milos, G., Spindler, A., Schnyder, U., & Fairburn, C. (2005). Instability of eating disorder diagnoses: Prospective study. *The British Journal of Psychiatry:*

The Journal of Mental Science, *187*, 573–578. https://doi.org/10.1192/bjp.187.6.573

Modesti, P. A., Reboldi, G., Cappuccio, F. P., Agyemang, C., Remuzzi, G., Rapi, S., … ESH Working Group on CV Risk in Low Resource Settings. (2016). panethnic differences in blood pressure in europe: A systematic review and meta-analysis. *PLOS ONE*, *11*(1), e0147601. https://doi.org/10.1371/journal.pone.0147601

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mondini, S., Mapelli, D., Vestri, A., Arcara, G., & Bisiacchi, P. (2011). *L'Esame Neuropsicologico Breve—2: Una batteria di test per lo screening neuropsicologico [Short neuropsychological exam 2: A battery of tests for neuropsychological screening].* Milano: Raffaello Cortina.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide: Statistical Analysis with Latent Variables (7th ed.)*. Los Angeles, CA: Authors.

National Institute for Health and Care Excellence. (2017). *Eating disorders: Recognition and treatment*. Retrieved from https://www.nice.org.uk/guidance/ng69

Nelson, H. E., Willison, J. (1991). *National Adult Reading Test (2nd ed.)* Windsor: NFER-Nelson.

Nikendei, C., Funiok, C., Pfüller, U., Zastrow, A., Aschenbrenner, S., Weisbrod, M., … Friederich, H. C. (2011). Memory performance in acute and

weight-restored anorexia nervosa patients. *Psychological Medicine*, *41*(4), 829–838. https://doi.org/10.1017/S0033291710001121

Norring, C. E. A., & Sohlberg, S. S. (1993). Outcome, recovery, relapse and mortality across six years in patients with clinical eating disorders. *Acta Psychiatrica Scandinavica*, *87*(6), 437–444. https://doi.org/10.1111/j.1600-0447.1993.tb03401.x

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., … Marder, S. R. (2008). The MATRICS Consensus Cognitive Battery, part 1: Test selection, reliability, and validity. *The American Journal of Psychiatry*, *165*(2), 203–213. https://doi.org/10.1176/appi.ajp.2007.07010042

Nunn, K., Frampton, I., Gordon, I., & Lask, B. (2008). The fault is not in her parents but in her insula--a neurobiological hypothesis of anorexia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *16*(5), 355–360. https://doi.org/10.1002/erv.890

Papadopoulos, F. C., Ekbom, A., Brandt, L., & Ekselius, L. (2009). Excess mortality causes of death and prognostic factors in anorexia nervosa. *The British Journal of Psychiatry*, *194*(1), 10–17. https://doi.org/10.1192/bjp.bp.108.054742

Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology*, *32*(1), 8–47. https://doi.org/10.1016/j.cedpsych.2006.10.003

Peat, C., Mitchell, J. E., Hoek, H. W., & Wonderlich, S. A. (2009). Validity and utility of subtyping anorexia nervosa. *International Journal of Eating Disorders*, *42*(7), 590–594. https://doi.org/10.1002/eat.20717

Pryor, T., Wiederman, M. W., & McGilley, B. (1996). Clinical correlates of anorexia nervosa subtypes. *The International Journal of Eating Disorders*, *19*(4), 371–379.

Rauch, S. L., & Savage, C. R. (1997). Neuroimaging and neuropsychology of the striatum. Bridging basic science and clinical practice. *The Psychiatric Clinics of North America*, *20*(4), 741–768.

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*(3), 271–276. https://doi.org/10.2466/pms.1958.8.3.271

Renwick, B., Musiat, P., Lose, A., DeJong, H., Broadbent, H., Kenyon, M., … Schmidt, U. (2015). Neuro- and social-cognitive clustering highlights distinct profiles in adults with anorexia nervosa. *The International Journal of Eating Disorders*, *48*(1), 26–34. https://doi.org/10.1002/eat.22366

Rey, A. (1964). *L'Examen Clinique en Psychologie [The Clinical Examination in Psychology]*. Presses Universitaires de France: Paris.

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): A factor analytic study of a large sample of normal elderly volunteers. *Dementia (Basel, Switzerland)*, *5*(5), 266–281.

Roberts, M. E., Tchanturia, K., Stahl, D., Southgate, L., & Treasure, J. (2007). A systematic review and meta-analysis of set-shifting ability in eating

disorders. *Psychological Medicine*, *37*(8), 1075–1084. https://doi.org/10.1017/S0033291707009877

Roberts, M. E., Tchanturia, K., & Treasure, J. L. (2016). Overlapping neurocognitive inefficiencies in anorexia nervosa: A preliminary investigation of women with both poor set-shifting and weak central coherence. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, *21*(4), 725–729. https://doi.org/10.1007/s40519-016-0302-8

Rose, M., Davis, J., Frampton, I., & Lask, B. (2011). The Ravello Profile: Development of a global standard neuropsychological assessment for young people with anorexia nervosa. *Clinical Child Psychology and Psychiatry*, *16*(2), 195–202. https://doi.org/10.1177/1359104511401191

Rose, M., Frampton, I., & Lask, B. (2012). A case series investigating distinct neuropsychological profiles in children and adolescents with anorexia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *20*(1), 32–38. https://doi.org/10.1002/erv.1095

Rose, M., Stedal, K., Reville, M. C., van Noort, B. M., Kappel, V., Frampton, I., … Lask, B. (2016). Similarities and differences of neuropsychological profiles in children and adolescents with anorexia nervosa and healthy controls using cluster and discriminant function analyses. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*. https://doi.org/10.1093/arclin/acw068

Sato, Y., Saito, N., Utsumi, A., Aizawa, E., Shoji, T., Izumiyama, M., … Fukudo, S. (2013). Neural basis of impaired cognitive flexibility in patients with anorexia nervosa. *PLOS ONE*, *8*(5), e61108. https://doi.org/10.1371/journal.pone.0061108

Shafran, R., & Somers, J. (1998). Treating adolescent obsessive-compulsive disorder: Applications of the cognitive theory. *Behaviour Research and Therapy*, *36*(1), 93–97.

Shallice, T., Broadbent, D. E., & Weiskrantz, L. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Sherman, B. J., Savage, C. R., Eddy, K. T., Blais, M. A., Deckersbach, T., Jackson, S. C., … Herzog, D. B. (2006). Strategic memory in adults with anorexia nervosa: are there similarities to obsessive compulsive spectrum disorders? *The International Journal of Eating Disorders*, *39*(6), 468–476. https://doi.org/10.1002/eat.20300

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*(6), 553–566. https://doi.org/10.1080/09638230500347541

Stern, R. A., & White, T. (2003). *Neuropsychological Assessment Battery*. Florida: PAR.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford: Oxford University Press.

Strober, M. (1981). The significance of bulimia in juvenile anorexia nervosa: An exploration of possible etiologic factors. *International Journal of Eating Disorders*, *1*(1), 28–43. https://doi.org/10.1002/1098-108X(198123)1:1<28::AID-EAT2260010104>3.0.CO;2-9

Stroop, J. R. (1935). Studies of interference in serial verbal reactions, *Journal of Experimental Psychology*, *18*, 643-662.

Tamiya, H., Ouchi, A., Chen, R., Miyazawa, S., Akimoto, Y., Kaneda, Y., & Sora, I. (2018). Neurocognitive impairments are more severe in the binge-eating/purging anorexia nervosa subtype than in the restricting subtype. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00138

Tchanturia, K., Morris, R. G., Anderluh, M. B., Collier, D. A., Nikolaou, V., & Treasure, J. (2004). Set shifting in anorexia nervosa: an examination before and after weight gain, in full recovery and relationship to childhood and adult OCPD traits. *Journal of Psychiatric Research*, *38*(5), 545–552. https://doi.org/10.1016/j.jpsychires.2004.03.001

Tchanturia, K., Morris, R. G., 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: EWD*, *7*(4), 312–315.

Tchanturia, Kate, Campbell, I. C., Morris, R., & Treasure, J. (2005). Neuropsychological studies in anorexia nervosa. *International Journal of Eating Disorders*, *37*(S1), S72–S76. https://doi.org/10.1002/eat.20119

Tchanturia, Kate, Davies, H., & Campbell, I. C. (2007). Cognitive remediation therapy for patients with anorexia nervosa: preliminary findings. *Annals of General Psychiatry*, *6*(1), 14. https://doi.org/10.1186/1744-859X-6-14

Treasure, J., Claudino, A. M., & Zucker, N. (2010). Eating disorders. *The Lancet*, *375*(9714), 583–593. https://doi.org/10.1016/S0140-6736(09)61748-7

Uznadze, D, N. (1966). *The psychology of set.* New York: Consultants' Bureau.

Vall, E., & Wade, T. D. (2015). Trail making task performance in inpatients with anorexia nervosa and bulimia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *23*(4), 304–311. https://doi.org/10.1002/erv.2364

Van Autreve, S., De Baene, W., Baeken, C., van Heeringen, K., Vancayseele, N., & Vervaet, M. (2016). Differential neural correlates of set-shifting in the bingeing–purging and restrictive subtypes of anorexia nervosa: An fMRI study. *European Eating Disorders Review*, *24*(4), 277–285. https://doi.org/10.1002/erv.2437

Van Autreve, S., De Baene, W., Baeken, C., van Heeringen, C., & Vervaet, M. (2013). Do restrictive and bingeing/purging subtypes of anorexia nervosa differ on central coherence and set shifting? *European Eating Disorders Review*, *21*(4), 308–314. https://doi.org/10.1002/erv.2233

Ward, A., Ramsay, R., Turnbull, S., Steele, M., Steele, H., & Treasure, J. (2001). Attachment in anorexia nervosa: A transgenerational perspective. *British Journal of Medical Psychology*, *74*(4), 497–505. https://doi.org/10.1348/000711201161145

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children–Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1981) *Wechsler Adult Intelligence Scale–Revised.* New York: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale–Revised*. New York: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale–Third Edition manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2010). *Wechsler Adult Intelligence Scale–Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Weigl, E., (1941). On the psychology of so-called process of abstraction, *Journal of Abnormal and Social Psychology*, *36*, 3 –33.

Weiß, R. H. (2006). *Grundintelligenztest—Revision (CFT-20-R)*. Göttingen: Hogrefe Ltd.

Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., Tugwell, P. (2001). The Newcastle-Ottawa Scale (NOS) for assessing the quality of non randomised studies in meta-analyses. Retrieved from *http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp*.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the Group Embedded Figures Test*. Menlo Park, CA: Mind Garden Inc.

Yano, M., Kawano, N., Tanaka, S., Kohmura, K., Katayama, H., Nishioka, K., & Ozaki, N. (2016). Dysfunction of response inhibition in eating disorders. *Journal of Clinical and Experimental Neuropsychology*, *38*(6), 700–708. https://doi.org/10.1080/13803395.2016.1151480

Zucker, N. L., Losh, M., Bulik, C. M., LaBar, K. S., Piven, J., & Pelphrey, K. A. (2007). Anorexia nervosa and autism spectrum disorders: Guided investigation of social cognitive endophenotypes. *Psychological Bulletin*, *133*(6), 976–1006. https://doi.org/10.1037/0033-2909.133.6.976

Appendices

Appendix A. Newcastle-Ottawa Scale – Cross Sectional

Selection:

1. Representativeness of the sample:
    a. Truly representative of the average in the target population. * (all subjects or random sampling)
    b. Somewhat representative of the average in the target group. * (non-random sampling)
    c. Selected group of users/convenience sample.
    d. No description of the derivation of the included subjects.
2. Sample size:
    a. Justified and satisfactory (including sample size calculation). *
    b. Not justified.
    c. No information provided
3. Non-respondents:
    a. Proportion of target sample recruited attains pre-specified target or basic summary of non-respondent characteristics in sampling frame recorded. *
    b. Unsatisfactory recruitment rate, no summary data on non-respondents.
    c. No information provided
4. Ascertainment of the exposure (risk factor):
    a. Vaccine records/vaccine registry/clinic registers/hospital records only. **
    b. Parental or personal recall and vaccine/hospital records. *
    c. Parental/personal recall only.

Comparability: (Maximum 2 stars)

1. Comparability of subjects in different outcome groups on the basis of design or analysis. Confounding factors controlled.
    a. Data/ results adjusted for relevant predictors/risk factors/confounders e.g. age, sex, time since vaccination, etc. **
    b. Data/results not adjusted for all relevant confounders/risk factors/information not provided.

Outcome:

1. Assessment of outcome:
    a. Independent blind assessment using objective validated laboratory methods. **
    b. Unblinded assessment using objective validated laboratory methods. **
    c. Used non-standard or non-validated laboratory methods with gold standard. *
    d. No description/non-standard laboratory methods used.
2. Statistical test:

a. Statistical test used to analyse the data clearly described, appropriate and measures of association presented including confidence intervals and probability level (p value). *
b. Statistical test not appropriate, not described or incomplete.

Cross-sectional Studies:

- Very Good Studies: 9-10 points
- Good Studies: 7-8 points
- Satisfactory Studies: 5-6 points
- Unsatisfactory Studies: 0 to 4 points

This scale has been adapted from the Newcastle-Ottawa Quality Assessment Scale for cohort studies to provide quality assessment of cross-sectional studies (Herzog, 2013).

Appendix B. Glossary

ASIT	Arrow-space Interference Task
BACS-SC	Brief Assessment of Cognition in Schizophrenia – Symbol Coding test
BIT	Brief Intelligence Test
BD	Block Design
BSAT	Brixton Spatial Anticipation Test
BVMT-R	Brief Visuospatial Memory Test – Revised
CANTAB	Cambridge Neuropsychological Test Automated Battery
CFAN	Category Fluency – Animal Naming
CFT-20-R	Culture Fair Test 20-Revised
COWAT	Controlled Oral Word Association Test
CPT-IP	Continuous Performance Test – Identical Pairs
CWI	Colour-Word Interference, d2TA = d2 Test of Attention
DKEFS	Delis-Kaplan Executive Functional System
IED	Intra Dimensional/Extra Dimensional Shift Task
IGT	Iowa Gambling Task
GEFT	Group Embedded Figures Test
HI	Haptic Illusions
HVLT-R	Hopkins Verbal Learning Test – Revised
LNS	Letter-Number Span
MCCB-J	Japanese MATRICS Consensus Cognition Battery
NAB-M	Neuropsychological Assessment Battery – Mazes
NART	National Adult Reading Test
NART-J	Japanese NART
MWIT	Memory With Interference Task
OA	Object Assembly
OAT	Object Alternation Test
RAVLT	Ray Auditory Verbal Learning Task
RCFT	Ray Complex Figure Test
RMF	Reading the Mind in Films Task
RWFT	Regensburg Word Fluency Test
SA	Slips-of-Action
SS	Spatial Span
SST	Stop Signal Task
STP	Stroop task
TMT	Trail Making Task
TSP	a task switching paradigm
TT	Tower Tests
VF	Verbal Fluency
WAIS	Wechsler Adult Intelligence Scale
WCST	Wisconsin Card Sorting Task
WISC-R	Wechsler Intelligence Scale for Children Revised
WMS-R	Wechsler Memory Scale Revised
WMS-III	Wechsler Memory Scale III
WST	Weigl’s Sorting Test.

Appendix C. Dissemination Statement

The study results will be disseminated locally at the Exeter Eating Disorder Research Group, chaired by Dr Ian Frampton, and to the 3rd year DClinPsy cohort on 10th June 2019. The literature review will be submitted for publication in the European Eating Disorders Review. Finally, study outcomes will be submitted for presentation at professional conferences.

Appendix D. European Eating Disorders Review Submission Guidelines

1. SUBMISSION

Authors should kindly note that submission implies that the content has not been published or submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium.

2. AIMS AND SCOPE

European Eating Disorders Review provides an international forum for disseminating cutting-edge theoretical and empirical research that significantly advances understanding of the relationship between Eating Disorders and Abnormal Eating/Weight conditions and well-being in humans.

Authors may submit original theoretical systematic reviews, methodological, or empirical research articles (7000 words or less) or short communications (3000 words or less). The aims of the journal are to offer a channel of communication between researchers, practitioners, administrators and policymakers who need to report and understand developments in the field of eating disorders.

3. MANUSCRIPT CATEGORIES AND REQUIREMENTS

Review articles: Systematic and meta-analytic review papers are welcomed if they critically review the available literature in a topic than will enhance clinical practice. Articles should have clear focus and enough number of studies should be available for a substantive review paper. Studies that only describe or list previous studies without a critical overview of the literature will not be considered.
Word Limit: 5,000 (excluding abstract, references, tables or figures).
Abstract: 200 words.
References: up to 100.
Figures/Tables: 5 maximum, but should be appropriate to the material covered. Additional tables might be included as supplementary information, if needed. Review articles must follow the PRISMA Guidelines. Authors may want to have a look at the review check lists that reviewers when assessing review articles.

Main Text File

The text file should be presented in the following order:

i. A short title that contains the major key words. The title should not contain abbreviations (see Wiley's best practice SEO tips);
ii. A short running title of less than 40 characters;
iii. The full names of the authors;
iv. The author's institutional affiliations where the work was conducted, with a footnote for the author’s present address if different from where the work was conducted;
v. The corresponding author’s contact email address and telephone number;
vi. Acknowledgments;
vii. Conflict of Interest statement (for all authors)
viii. Names and grant numbers of any sources of funding or support in the form of grants, equipment, drugs etc.

Acknowledgments

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.

Conflict of Interest Statement

Authors will be asked to provide a conflict of interest statement during the submission process. For details on what to include in this section, see the section 'Conflict of Interest' in the Editorial Policies and Ethical Considerations section below. Submitting authors should ensure they liaise with all co-authors to confirm agreement with the final statement.

Main Text File

As papers are double-blind peer reviewed, the main text file should not include any information that might identify the authors.

The main text file should be presented in the following order:

i. Title, abstract, highlights and key words;
ii. Main text;
iii. References;
iv. Tables (each table complete with title and footnotes);
v. Figure legends;
vi. Appendices (if relevant).

Figures and supporting information should be supplied as separate files.

Abstract

All manuscripts should contain an abstract of up to 200 words. An **abstract** is a concise summary of the whole paper, not just the conclusions, and is understandable without reference to the rest of the paper. It should contain no citation to other published work. It must be structured, under the sub-headings: Objective; Method; Results; Conclusions.

Highlights

Highlights are mandatory for European Eating Disorders Review. These should appear as three bullet points that convey the core findings of the article.

Keywords

Include up to five **keywords** that describe your paper for indexing purposes.

References

References should be prepared according to the Publication Manual of the American Psychological Association (6th edition). This means in text citations should follow the author-date method whereby the author's last name and the year of publication for the

source should appear in the text, for example, (Jones, 1998). The complete reference list should appear alphabetically by name at the end of the paper. A sample of the most common entries in reference lists appears below. Please note that a DOI should be provided for all references where available. For more information about APA referencing style, please refer to the APA FAQ. Please note that for journal articles, issue numbers are not included unless each issue in the volume begins with page one.

Figure Legends

Legends should be concise but comprehensive – the figure and its legend must be understandable without reference to the text. Include definitions of any symbols used and define/explain all abbreviations and units of measurement.

Figures

Although authors are encouraged to send the highest-quality figures possible, for peer-review purposes, a wide variety of formats, sizes, and resolutions are accepted. Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post-acceptance figure requirements.

Additional Files

Appendices

Appendices will be published after the references. For submission they should be supplied as separate files but referred to in the text.

Supporting Information

Supporting information is information that is not essential to the article, but provides greater depth and background. It is hosted online and appears without editing or typesetting. It may include tables, figures, videos, datasets, etc. Click here for Wiley's FAQs on supporting information.

Note: if data, scripts, or other artefacts used to generate the analyses presented in the paper are available via a publicly available data repository, authors should include a reference to the location of the material within their paper.

If a manuscript describes a new approach and/or technological approach, authors are encouraged to include a small demo video – no more than 60 seconds long.

General Style Points

The following points provide general advice on formatting and style.

- **Language:** The language of the journal is English.
- **Abbreviations:** In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Initially, use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only. Distinction should be made between capital and lower case letters, between the letter O and zero, between the letter I and number one and prime, between k and kappa.

- **Units of measurement:** Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website at www.bipm.fr for more information about SI units.
- **Numbers:** numbers under 10 are spelt out, except for: measurements with a unit (8mmol/l); age (6 weeks old), or lists with other numbers (11 dogs, 9 cats, 4 gerbils).
- **Trade Names:** Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name and the name and location of the manufacturer in parentheses.

5. EDITORIAL POLICIES AND ETHICAL CONSIDERATIONS

Data Storage and Documentation

European Eating Disorders Review encourages authors to share the data and other artefacts supporting the results in the paper by archiving it in an appropriate public repository. Authors should include a data accessibility statement, including a link to the repository they have used, in order that this statement can be published alongside their paper.

Authors can consult the global registry of research data repositories re3data.org to help them identify registered and certified repositories relevant to their subject areas.

Data Citation In recognition of the significance of data as an output of research effort, Wiley has endorsed the FORCE11 Data Citation Principles and is implementing a mandatory data citation policy. Journal policies should require data to be cited in the same way as article, book, and web citations and authors are required to include data citations as part of their reference list. Data citation is appropriate for data held within institutional, subject focused, or more general data repositories. It is not intended to take the place of community standards such as in-line citation of GenBank accession codes. When citing or making claims based on data, authors must refer to the data at the relevant place in the manuscript text and in addition provide a formal citation in the reference list. We recommend the format proposed by the Joint Declaration of Data Citation Principles

Human Studies and Subjects

For manuscripts reporting medical studies that involve human participants, a statement identifying the ethics committee that approved the study and confirmation that the study conforms to recognized standards is required, for example: Declaration of Helsinki; US Federal Policy for the Protection of Human Subjects; or European Medicines Agency Guidelines for Good Clinical Practice. It should also state clearly in the text that all persons gave their informed consent prior to their inclusion in the study.

Patient anonymity should be preserved. Photographs need to be cropped sufficiently to prevent human subjects being recognized (or an eye bar should be used). Images and information from individual participants will only be published where the authors have obtained the individual's free prior informed consent. Authors do not need to provide a copy of the consent form to the publisher; however, in signing the author license to

publish, authors are required to confirm that consent has been obtained. Wiley has a standard patient consent form available for use.

Research Reporting Guidelines

Accurate and complete reporting enables readers to fully appraise research, replicate it, and use it. Authors are expected to adhere to the following research reporting standards.

- CONSORT checklist for reports of randomised trials and cluster randomised trials
- TREND checklist for non-randomised controlled trials
- PRISMA checklist for systematic reviews and meta-analyses • STROBE checklist for observational research
- COREQ checklist for qualitative studies
- SQUIRE checklist for quality improvement

See the EQUATOR Network for other study types.

Conflict of Interest

The journal requires that all authors disclose any potential sources of conflict of interest. Any interest or relationship, financial or otherwise that might be perceived as influencing an author's objectivity is considered a potential source of conflict of interest. These must be disclosed when directly relevant or directly related to the work that the authors describe in their manuscript. Potential sources of conflict of interest include, but are not limited to: patent or stock ownership, membership of a company board of directors, membership of an advisory board or committee for a company, and consultancy for or receipt of speaker's fees from a company. The existence of a conflict of interest does not preclude publication. If the authors have no conflict of interest to declare, they must also state this at submission. It is the responsibility of the corresponding author to review this policy with all authors and collectively to disclose with the submission ALL pertinent commercial and other relationships.

Funding

Authors should list all funding sources in the Acknowledgments section. Authors are responsible for the accuracy of their funder designation. If in doubt, please check the Open Funder Registry for the correct nomenclature:
https://www.crossref.org/services/funder-registry/

The list of authors should accurately illustrate who contributed to the work and how. All those listed as authors should qualify for authorship according to the following criteria:

1. Have made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; and
2. Been involved in drafting the manuscript or revising it critically for important intellectual content; and
3. Given final approval of the version to be published. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content; and
4. Agreed to be accountable for all aspects of the work in ensuring that questions related

to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section (for example, to recognize contributions from people who provided technical help, collation of data, writing assistance, acquisition of funding, or a department chairperson who provided general support). Prior to submitting the article all authors should agree on the order in which their names will be listed in the manuscript.

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

EMPIRICAL PAPER

Neuropsychological Profiles in Anorexia Nervosa and Their Demographic and Clinical Correlates

Trainee Name:	**Mr Alexander Charles Lightley Drake**
Primary Research Supervisor:	**Dr Ian Frampton** Consultant Clinical Psychologist
Secondary Research Supervisor:	**Dr Anke Karl** Senior Lecturer
Target Journal:	International Journal of Eating Disorders
Word Count:	8931 words (excluding abstract, table of contents, list of figures, references, footnotes, appendices)

Submitted in partial fulfilment of requirements for the Doctorate Degree in Clinical Psychology, University of Exeter

Acknowledgements

I would like to thank Drs Anke Karl and Ian Frampton for their support throughout the project. Thanks too to Drs Ian McLaren and Roland Baddeley, who provided timely help with my efforts to learn the statistics and programming required to implement a neural network analysis. Finally, I would like to thank my wife, Alice, for her constant love and patience.

Abstract

Objective: Treatment outcomes for anorexia nervosa (AN) remain unsatisfactory. Substantial research has investigated the neuropsychological effects of AN, often with mixed results. One explanation for the inconsistencies is that there exist several distinct neuropsychological profiles within AN. Profiles have been reported, though not associated with clinical or demographic variables, limiting their utility. Suboptimal statistical techniques may undermine these findings. **Method**: An existing dataset of healthy controls (HCs) and AN patients (n = 423) was subjected to secondary analysis using latent profile analysis and a neural network to investigate latent profiles and the existence of non-linear neuropsychological structure. Profiles were compared with respect to demographic and clinical variables. **Results**: The latent profile analysis revealed five AN neuropsychological profiles. Patients in a globally neuropsychologically impaired profile were older than those in a high-average with high verbal profile and weighed less than those in an average performance profile. A non-linear neural network failed to outperform a linear neural network on a diagnosis classification task. **Discussion**: The five-profile solution extended the neuropsychological groups previously found in the literature. This study is the first to successfully associate latent neuropsychological profile to clinically meaningful variables, though the profile in which differences were observed was tiny (7% of patients). None of the discovered profiles differed in terms of anxiety, undermining support for the noradrenergic hypothesis of AN. The failure of the non-linear neural network to outperform the linear network indicates that AN neuropsychological ability does not contain significant non-linearity, indicating that conventional statistical techniques can model them.

Introduction

Anorexia nervosa (AN), an eating disorder of self-starvation that is poorly understood and managed. Psychopharmacological interventions are ineffective. Cognitive-behavioural and family therapy are recommended (National Institute for Health and Care Excellence, 2017) and innovative approaches such as cognitive remediation therapy, used within AN to promote greater cognitive flexibility, are gaining traction (Tchanturia, Lloyd, & Lang, 2013). For nearly thirty years, researchers have been trying to understand the neuropsychological impact of AN (Touyz, Beumont, & Johnstone, 1986), with many differences between patients with AN and healthy controls (HCs) observed, though neuropsychological heterogeneity exists (Rose, Frampton, & Lask, 2012), implying that AN may comprise distinct neuropsychological profiles.

Neuropsychological Functioning

Differences in intelligence, central coherence, executive functioning, memory, visuospatial processing, and memory have all been observed between HCs and AN patients.

Intelligence. Research into the IQ of AN patients has produced varied results. Some authors report reduced intelligence among AN patients (Weider, Indredavik, Lydersen, & Hestad, 2014), meta-analysis indicates a significant moderate difference is observable between AN patients and HCs (Lopez, Stahl, & Tchanturia, 2010), and recent research using a matched case-control sample of 188 found no difference between the IQs of AN patients compared with HCs (Telléus et al., 2015).

Visuospatial processing. Visuospatial processing enables the visual perception and manipulation of objects in the mind and environment (Rauch &

Savage, 1997). A number of studies have reported deficits in AN visuospatial processing (Jones, Duncan, Brouwers, & Mirsky, 1991; Kim, Lim, & Treasure, 2011; Lopez, Tchanturia, Stahl, & Treasure, 2009), while others found AN performance equivalent to HCs (Castro-Fornieles et al., 2009; Danner et al., 2012; Stedal, Rose, Frampton, Landrø, & Lask, 2012). Visuospatial research in AN is hampered by the difficulty in discriminating between impairment in visuospatial functioning versus impairment in central coherence, as visuospatial tasks invariably involve complex visual stimuli that require central coherence to process properly.

Central coherence. The ability to shift focus from the detail to gestalt of incoming stimuli is known as central coherence. Weak central coherence characterises autism (Happé & Frith, 2006) and research increasingly suggests similar impairments may affect AN patients (Lena, Fiocco, & Leyenaar, 2004; Lopez et al., 2008). Central coherence deficits may persist after weight restoration (Lopez et al., 2009) and can be observed across a range of neuropsychological tests (Harrison, Tchanturia, & Treasure, 2011; Lopez et al., 2008; Southgate, Tchanturia, & Treasure, 2005). A recent meta-analysis described AN central coherence weakness as moderate-to-large (Lang, Lopez, Stahl, Tchanturia, & Treasure, 2014).

Visuospatial memory. A mixed picture exists regarding visuospatial memory deficits. Patients with AN have visuospatial memory and spatial recognition impairments (Fowler et al., 2006; Key, O'Brien, Gordon, Christie, & Lask, 2006), though do not differ with HCs on short-term and working memory nor recognition (Fowler et al., 2006). Compared to HCs, AN patients performed poorly on the Rey Copy and Rey Recall tests of immediate and delayed visual recall (Sherman et al., 2006).

Verbal memory. Evidence of AN verbal memory impairments is inconsistent. A number of studies report no difference between AN patients and HCs on verbal short-term memory (e.g. Kemps et al., 2006; Key et al., 2006; Szmukler et al., 1992), while a small number report AN patients performing less well than HCs (Castro-Fornieles et al., 2009; Green, Elliman, Wakeling, & Rogers, 1996). Evidence of verbal working memory deficits is scarce. The largest verbal working memory study found no difference between AN patients and HCs (Lao-Kaim, Giampietro, Williams, Simmons, & Tchanturia, 2014).

Verbal functioning. Early research suggested, compared to HCs, AN patients display modest deficits in verbal performance (Jones et al., 1991), though others report no difference (Steinglass, Walsh, & Stern, 2006; Tchanturia et al., 2004). A meta-analysis of verbal functioning findings actually reported that AN patients possessed better word-generation ability than HCs (Stedal, Frampton, Landrø, & Lask, 2012).

Executive functioning. Goal-directed behaviour depends on cognitive abilities such as memory, inhibition, set-shifting, problem-solving, and planning (Diamond, 2013). In sustained self-starvation, AN patients demonstrate extraordinary inhibitory capacity; yet, they frequently struggle to impair intrusive thoughts about body shape and weight (Shafran & Somers, 1998). Set-shifting deficits, impaired ability to switch attention from one stimulus or task to another in response to an environmental change, have been observed in AN patients, recovered AN patients, and the healthy sisters of patients with AN (Roberts, Tchanturia, Stahl, Southgate, & Treasure, 2007; Tchanturia, Campbell, Morris, & Treasure, 2005).

Despite the positive neuropsychological findings, controversy remains. Two comprehensive literature reviews concluded there was no evidence of neuropsychological deficits in AN (Duchesne et al., 2004; Lena et al., 2004). More recently, consensus developed that executive functioning, visuospatial processing, and central coherence were reliability impaired (Lopez et al., 2008; Roberts et al., 2007; Steinglass & Walsh, 2006), although others maintain that inconsistencies remain unresolved (Brewerton, Frampton & Lask, 2009). One explanation may be that neuropsychological deficits in AN are not uniform: there may exist several profiles of neuropsychological dysfunction.

Neuropsychological profiles

An early case series paper by Rose et al. (2012) selectively sampled from a larger dataset to demonstrate the heterogeneity observable within AN neuropsychological performance. They showed that the variance in executive functioning was much lower ($0.06 < z < 1.50$) than in visuospatial ($-2.45 < z < 1.35$) or central coherence ($-3.77 < z < 0.50$) performance.

Harrison, Tchanturia, Naumann, & Treasure (2012) compared set-shifting, central coherence, and theory of mind abilities between patients with AN and healthy controls using principal components analysis. AN patients scored highly on a fragmented, perseverative, and rigid cognitive style component and a socio-emotional difficulties component. Recovered patients did not differ from current patients. Patients with the most extreme neuropsychological deficits (defined as in the lowest 90th percentile) were more symptomatic than chronic AN patients, suggesting neurocognitive deficits play a role in maintaining AN symptoms and may underlie them (Schmidt & Treasure, 2006). Harrison et al. did not perform separate principal component analyses for

the patients and control groups, which may have obscured any AN-specific groups present in the data.

Other researchers have identified neuropsychological profiles of patients with AN. From a battery of set-shifting, central coherence and social-cognition tests administered to AN patients, three neuropsychological clusters were identified (Renwick et al., 2015). Patients within the three clusters – average to high-average neurocognitive and social-emotional performance, neuropsychological strengths and weaknesses, and poor overall performance – did not differ in symptoms, comorbidity, service utilisation, or treatment adherence, suggesting the clusters may not be clinically meaningful. Renwick et al. did not include a matched control group, so it is possible sampling error produced the results.

Further evidence of neuropsychological AN profiles has been reported by Rose et al. (2016), who found three AN clusters: neuropsychologically low-average to average, discrepant verbal and visuospatial ability, and neuropsychologically average to high-average with strong verbal ability. As the clusters were equivalent on clinical variables such as disordered cognitions, anxiety, depression, obsessionality, and BMI, they may lack clinical utility. Unexpectedly, central coherence was greater in AN patients than the controls, contrary to existing evidence (Lang & Tchanturia, 2014).

Not all researchers have found evidence for neuropsychological profiles in AN. Bentz et al. (2017) compared set-shifting, central coherence, processing speed, working memory, sustained attention, verbal memory, and verbal abstraction between first-episode AN patients, recovered AN patients, and healthy controls. They found no difference in any of the neurocognitive abilities between the three groups.

The Noradrenergic Model

Findings of AN neuropsychological deficits caused speculation that the disorder may be neuroanatomical. Researchers have questioned whether dysfunction in the insula lobe of the brain may underlie the disparate symptoms of AN (Nunn, Frampton, Gordon, & Lask, 2008). The insula lobe is connected to frontal, temporal, and parietal lobes; the limbic system (amygdala, hippocampus, thalamus and hypothalamus); nucleus accumbens; and the striatum (Shelley & Trimble, 2004). The insula is thought to be a cortical hub that balances the competing needs for internal homeostasis and adaptation to the environment (Mesulam & Mufson, 1988). The noradrenergic model suggests AN develops due to a threat response that cannot be integrated by the insula with interoceptive information received from temporal somatosensory cortex or satiety information from basal ganglia structures, causing anxiety and restricted eating (Nunn et al., 2008). The onset of starvation behaviour is highly negatively reinforcing because it reduces the availability of precursors required to synthesise noradrenaline, which reduces insula activity and alleviates the anxiety. Eating becomes positively punishing because it provides the nutrients required to reactivate the insula, thus increasing anxiety (Nunn, Frampton, & Lask, 2012). While the insular lobe is thought to be involved in many bodily functions, it is unclear how insula dysfunction might directly affect a person's neuropsychological performance. Human cognitive performance varies considerably in the general population (Wechsler, 2010). It is possible that insular dysfunction may introduce additional variability that may be detectable as a distinct neuropsychological profile, though it is hard to predict what that profile might look like. Alternatively, it may be that insula dysfunction produces no direct effects on neuropsychological performance.

Suboptimal statistical procedures may explain the failure to link neuropsychological profiles with clinical and demographic factors. Principal components analysis, as used by Harrison et al. (2012), is a data reduction technique, not a method for identifying unobserved groups within data. Cluster analysis, as used by Renwick et al. (2015) and Rose et al. (2016), identifies groups based on similarities and differences within the data – but is subjective and lacks a goodness of fit metric (Ketchen & Shook, 1996). Latent profile analysis (LPA), a type of structural equation modelling, is a superior method that solves both problems. LPA provides test statistics for measuring model fit and because models are specified *a priori*, reduces the subjectivity inherent in cluster analysis (Berlin, Williams, & Parra, 2014).

Another approach that may help to understand the extant neuropsychological findings is machine learning. Machine learning neural networks (NN) have two distinct advantages over null-hypothesis significant testing approaches. NNs are atheoretical (unlike LPA a model is not specified in advance) and can model non-linear data (Bishop, 2006). The second advantage may be crucial for AN research, as the failure to link neuropsychological profiles to clinical and demographic variables may be due to the relationships between the data being non-linear.

If neuropsychological profiles of AN can reliably be identified and associated with clinical factors of interest such as BMI, cognitions, treatment adherence and response, we can improve our understanding of the disorder. We could develop neuropsychological profile-informed treatment pathways and accurate prognosis may be improved. Clinically-relevant neurocognitive profiles could form the basis of brain-directed treatments (e.g. cognitive-remediation

therapy). Linking neuropsychological function to anxiety would also provide support for the noradrenergic hypothesis of AN.

Research Questions

The present study aimed to answer the following research questions:

1) Can latent profiles of neuropsychological performance be detected among AN patients?

2) Do clinical and demographic variables differ between latent neuropsychological profiles?

3) Does the neuropsychological data contain significant non-linear structure? If so, can a non-linear model predict patient clinical and demographic variables?

It is predicted that the neuropsychological data will be non-linear, latent profiles will exist and that anxiety will differ between latent neuropsychological profiles.

Method

Design

The study will compare cross-sectional neuropsychological performance within AN patients and HCs.

Sample

The sample comprised 253 female children and adolescents with AN determined by DSM-IV (American Psychiatric Association, 1994) or Great Ormond Street hospital diagnostic criteria (Bryant-Waugh, 2000) and 170 healthy female controls (HCs). Participants were aged inclusively between 9

and 18 years (*M* = 15.6, *SD* = 1.8) and were recruited as part of the Ravello trial (Rose et al., 2016).

Measures

The dataset comprised age, BMI, BMI centile (Cole, Freeman, & Preece, 1995), and clinical and neuropsychological variables. For more detail on the Ravello profile, see Rose et al. (2011) and Rose et al. (2016).

Wechsler Abbreviated Scale of Intelligence (WASI). The WASI (Wechsler, 1999) is an abbreviated test of intelligence persons aged 6-69 years old. Matrix Reasoning and Vocabulary subtests were administered and used to prorate an IQ score.

Brixton Spatial Anticipation Test (BSAT). The BSAT (Burgess & Shallice, 1997) is a neuropsychological test that assesses the planning and disinhibition.

Delis-Kaplan Executive Functioning System (D-KEFS). The D-KEFS (Delis, Kaplan & Kramer, 2001) is a battery test of executive functioning. Colour-Word Interference (CWI) conditions 3 and 4; Verbal Fluency (VC) conditions 1, 2, and 3; Trial Making Test (TMT) condition 4; and Tower Test (TT) were used.

Rey Complex Figure Test (RCFT). The RCFT (Meyers & Meyers, 1996) assesses central coherence and visual perception**.** Immediate Recall (IR), Delayed Recall (DR), Recognition Trial (RT) were administered and used to compute the Central Coherence Index (CCI; Booth, 2006).

Eating Disorder Examination (EDE). The EDE (Cooper & Fairburn, 1987) is a semi-structured interview that measures psychological components

of eating disorders. It has four subscales – restraint, eating concern, shape concern, and weight concern – each with a maximum score of six.

State-Trait Anxiety Inventory (STAI). The STAI (Spielberger, Gorsuch, Lushene, Vagg & Jacobs, 1983) is a 40-item self-report questionnaire that measures state and trait anxiety.

Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). The Y-BOCS (Goodman et al., 1989) is a 10-item self-report questionnaire that measures obsessive and compulsive thoughts in the previous week.

Children's Obsessive-Compulsive Inventory (CHOCI). The CHOCI (Shafran et al., 2003) is a 21-item self-report questionnaire that measures compulsions and obsessions and their associated impairments in 7-17-year-olds.

Beck Depression Inventory (BDI). The BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) is a 21-item self-report questionnaire that measures symptoms of depression in those aged 13 or more.

Child Depression Inventory (CDI). The CDI (Kovacs, 1992) is a 27-item self-report questionnaire that measures symptoms of depression in 7-17-year-olds.

Procedure

The fully-anonymised Ravello dataset was subjected to secondary analysis.

Data Analysis

The research questions will be answered using LPA and NN analysis. Each procedure is described separately.

Latent profile analysis. LPA is a technique for estimating case membership of a latent, unmeasured categorical variable from observed continuous variables (Muthén & Muthén, 1998-2012). A conceptual description of LPA is provided in Appendix F.

The LPA was performed with MPlus 7.4 (Muthén & Muthén, 1998-2012) using maximum likelihood estimation, the most commonly used estimation method (Pastor, Barron, Miller, & Davis, 2007). Model selection was based on an iterative process using model fit statistics described by (Nylund, Asparouhov, & Muthén, 2007): the Bayesian information criterion (BIC; Schwarz, 1978), bootstrapped likelihood ratio test (BLRT; McLachlan & Peel, 2000), and entropy. The BIC is a likelihood-based measure of model fit that penalises for complexity (i.e. many model parameters); the BLRT is another likelihood-based test statistic, used to compare the fit of two nested models. It produces a *p*-value to assess whether adding a parameter *k* when compared with the *k*–1 model improves model fit. A two-profile model was created, then compared with the fit statistics of a three-factor model, a four-factor model, and so on.

The best fitting model will show the lowest BIC, a BLRT close to zero, a significant BLRT *p*-value, and profiles with high posterior probabilities (i.e. the probability that an individual belongs to the assigned profile is high and low for all other profiles). Individuals should be assigned to their profile with high probability ($p > 0.70$) and no profile should contain fewer than 5% of the total sample (Stanley, Kellermanns, & Zellweger, 2017).

The MPlus code used can be found in Appendix G. The MPlus STARTS values, which indicate how many initial stage random sets of starting values to generate and the number of final stage optimisations, were increased from the default 20-4 to 100-20 to produce the final replicable model.

Once the optimal model was found, neuropsychological test performance was compared between the profiles using multiple ANOVAs, with profile membership as the independent variable and the neuropsychological test scores as dependent variables. The Bonferonni correction was used for post hoc comparisons. For variables that did not meet the assumption of homogeneity of variance, one-way Brown-Forsythe ANOVAs with Games-Howell post hoc comparisons were used. Effect sizes were calculated using partial eta squared (η_p^2) and defined as small = 0.01, medium = 0.06, and large = 0.14 (Cohen, 1988). Profiles were then compared with clinical and demographic variables age, BMI, eating disorder pathology, anxiety, depression, and obsessionality using multiple one-way ANOVAs with Bonferronni's correction or a Brown-Forsythe ANOVA with Games-Howell post hoc test.

Neural networks. The data was analysed for non-linearity using NN: a method for building classifier or regressor functions based on iterative supervised learning. A conceptual description of NNs is provided in Appendix H.

The NN was created using Python v3.6.7 using the following packages: Keras 2.2.4, Matplotlib 3.0.1, NumPy 1.15.4, SciKit-Learn v0.20.0, Pandas 0.23.4, and TensorFlow 1.12.0. TensorFlow is a machine learning package developed by Google Brain, Keras is an application programming interface for using TensorFlow, ScitKit-Learn provides additional machine learning tools, NumPy is a scientific computing package, Matplotlib is a plotting library, and Pandas is a data analysis library. The code used in the NN analysis can be found in Appendix I.

A classifier NN was built to predict diagnostic status (AN vs. HC) from the 14 neuropsychological measures. The batch and epoch parameters used were determined by iterative testing. Batch sets the number of training examples that are inputted into the NN algorithm before the weights are updated. Epoch determines how many times the entire training data is inputted into the algorithm. Iterative testing was performed with 10-fold cross-validation on a training set comprising 80% of the total dataset. To minimise overfitting, a dropout rate of 0.1 was set across all neurons within the NN. Dropout randomly deactivates 10% of the neurons in each layer of the NN during each batch of training, effectively training a sub-sample of the network (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014). The RandomUniform initialiser was used to randomly generate weights with a uniform distribution in the range $-0.05 < w < 0.05$. The loss function used was binary cross entropy and the optimiser was Adam (Kingma & Ba, 2014).

The number of first layer neurons (7, 9, 14, 27), layers (1, 2, 3, 4), subsequent layer neurons (27, 14, 9, 7, 6, 5, 4, 3), epochs (100, 250, 500, 1000), and batch size (1, 2, 5, 10, 25, 32) were each tested using a leaky rectified linear unit (ReLU) function. The ReLU function, $f(x) = max(0, x)$, allows the network to model non-linear relationships within the data. The leaky ReLU function

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases} \quad (1)$$

does likewise, while avoiding the problem of dead neurons reducing learning within parts of the network (Maas, Hannun & Ng, 2013). The sigmoid activation function

$$f(x) = \frac{e^x}{e^x + 1} \quad (2)$$

was used in the output layer neuron to convert its input into a probability value ranging from 0 to 1, which, if greater than 0.5, predicts AN diagnosis. If the non-linear NN is superior to the linear network at predicting AN diagnosis, it will then be used to predict other variables from the neuropsychological data.

The optimised non-linear model was compared against the same model using the identity function in place of the ReLU function. The identity function, $f(x) = x$, limits the network to modelling only linear relationships. A modified independent-samples t-test, the 5x2cv, was used. The 5x2cv test was designed for comparing computational models using k-fold cross-classification. It is superior to alternatives, for example, the resampled paired t-test (based on Student's t-test), as it does not violate the assumptions of independence and normality, and has been shown to produce fewer type 1 errors (Dietterich, 1998). A full description of the 5x2cvtest can be found in Appendix J.

Power

Dziak, Lanza, & Tan (2014) estimated the sample sizes required to reliably find clusters with a large effect size (*w* = 0.5, comparable to *d* = 0.5) from 15 input variables was 297 using latent class analysis. They estimated the number of samples required to identify five classes with a large effect size was 462 (*w* = 0.4).

There is no empirically-based guidance for NN sample size requirements; however, a local professor with years of experience using NNs in psychology advised that the present sample was sufficient (I. McLaren, personal communication, May 19, 2017).

Ethical Approval

The study was based on secondary analyses of a fully anonymised dataset generated from an earlier study that had received appropriate research ethical approval in Norway and the UK (Appendix K).

Results

Anorexia Nervosa Profiles

A five-profile model produced the optimum LPA solution to the AN neuropsychological test data. The five-profile model provided the lowest BIC and BLRT (Table 1). The five-profile solution comprised profiles each with at least 5% of the sample. All average posterior probabilities were above 80%, indicating high certainty about case classification (Table 2).

Table 1.

Anorexia Nervosa Latent Profile Analysis Models

	Fit statistic			Profile membership distribution				
Profile	BIC	BLRT	Entropy	1	2	3	4	5
1	9885.03	NA	NA	254				
2	9469.86	–4865.00*	0.856	105	149			
3	9363.98	–4521.41*	0.870	24	103	127		
4	9308.05	–4451.91*	0.844	24	78	98	54	
5	9274.09	–4393.40*	0.854	18	41	69	76	50

Note. BIC = Bayesian information criterion, BLRT = bootstrapped likelihood ratio test, * = *p*-value < 0.05 indicating the k-profile model is superior than the *k*–1-profile model.

Table 2.

Five-profile Anorexia Nervosa Model Posteriors

Profile	*n*	%	1	2	3	4	5
1	18	7	**0.96**	0.01	0.04	0.00	0.00
2	41	16	0.00	**0.90**	0.07	0.04	0.00
3	69	27	0.00	0.02	**0.90**	0.06	0.01

4	76	30	0.00	0.01	0.06	**0.90**	0.03
5	50	20	0.00	0.00	0.02	0.05	**0.93**

Note. Bold values report the average posterior probability associated with the profile to which each case was assigned.

The between-subjects ANOVAs conducted on the neuropsychological test scores with LPA profile as the grouping variable indicated significant differences between every test ($p < 0.001$) except the BSAT ($p = 0.105$). The effect size was large for every variable ($\eta_p^2 > 0.14$), except the CCI ($\eta_p^2 = 0.10$), where it was moderate, and TT and BSAT ($0.03 < \eta_p^2 < 0.05$), where it was small (Table 3). A plot of the neuropsychological performance of each profile is presented in Figure 1.

The results of the between-profile comparisons were used to name the five AN profiles. Profile 1 ($n = 18$) contained average scores on central coherence and the Tower Test ($–0.5 < z < 0.5$), low-average ($–1 < z < 0$) IQ, recognition, verbal fluency, and BSAT scores, and very low ($–1 < z < –2$) immediate recall, delayed recall, color-word conditions 3 and 4 and the trails task. Profile 1 was named *globally impaired*. Profile 2 ($n = 41$) displayed severely impaired ($z = – 1.99$) immediate and delayed recall ($–2.10 < z < – 1.99$), low-average recognition ($z = –0.82$) and average ($–1 < z < 1$) performance in all other tests. Profile 2 was named *memory impaired*. Profile 3 ($n = 69$) scored in the average range ($–1 < z < 1$) on all tests and was named *average performance*. Profile 4 ($n = 76$) was also average ($0 < z < 1$), except in the tests of verbal fluency, where performance was above average ($z > 1$), and in visuospatial tests, where performance was low-average ($–0.83 < z < –0.90$). Profile 4 was named *visuo-verbal discrepancy*. Finally, profile 5 ($n = 50$) was named *high-average with strong verbal* as most neuropsychological scores

Table 3.

Neuropsychological Performance Between Anorexia Nervosa Profiles

	Profile										Test statistic			
	1 (*n* = 18)		2 (*n* = 41)		3 (*n* = 69)		4 (*n* = 76)		5 (*n* = 49)					
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	*p*	η_p^2	Post hoc
IQ	–0.93	0.60	–0.59	0.88	–0.37	0.95	0.27	0.86	0.95	0.93	28.57	**<0.001**	0.32	4, 5 > 3, 2, 1
IR	–1.34	0.86	–1.99	0.79	–0.42	0.68	–0.84	0.69	0.68	0.76	82.92	**<0.001**	0.57	All ≠ all, except 4 = 1
DR	–1.51	0.81	–2.10	0.89	–0.39	0.71	–0.90	0.72	0.70	0.68	88.50	**<0.001**	0.59	3, 4 > 1, 2; 5 > all
RT	–0.80	1.36	–0.82	1.28	–0.37	1.06	–0.83	0.83	0.41	0.75	11.33†	**<0.001**	0.18	5 > 1, 2, 3, 4; 4 > 3
CCI	0.37	0.70	–0.06	0.94	0.49	0.84	0.37	0.89	0.86	0.78	6.70	**<0.001**	0.10	3, 5 > 2; 5 > 4
VC1	–0.26	0.86	0.32	0.93	0.35	0.95	1.64	0.87	1.55	1.07	33.97	**<0.001**	0.35	4, 5 > 1, 2, 3
VC2	–0.16	0.95	0.50	1.01	0.90	0.71	1.74	0.78	1.64	0.85	33.15	**<0.001**	0.35	4, 5 > 1, 2, 3; 3 > 1
VC3	–0.52	1.06	0.70	0.96	0.68	0.73	2.05	0.68	1.83	0.89	49.57†	**<0.001**	0.49	4, 5 > 1, 2, 3; 2, 3 > 1
VC4	–0.58	0.97	0.91	0.81	0.66	0.74	1.97	0.70	1.83	0.73	63.20	**<0.001**	0.51	4, 5 > 1, 2, 3; 2, 3 > 1
CW3	–1.54	1.26	0.53	0.62	0.24	0.78	0.65	0.60	0.74	0.57	29.16†	**<0.001**	0.40	4, 5 > 3, 1; 3, 2 > 1
CW4	–1.95	1.30	0.16	0.95	0.42	1.02	0.33	0.80	0.63	0.79	23.99†	**<0.001**	0.31	5, 4, 3, 2 > 1
TT	–0.25	0.78	–0.04	0.67	0.08	0.78	–0.09	0.64	0.27	0.63	3.06	**0.017**	0.05	5 > 4
TMT	–1.47	1.06	–0.24	0.81	–0.34	0.96	0.04	0.66	0.37	0.84	15.95†	**<0.001**	0.22	5 > 3, 2, 1; 4 > 3; 4, 3, 2 > 1
BSAT	–0.57	1.30	–0.50	1.16	–0.25	1.02	–0.26	1.25	0.09	1.07	1.94	0.105	0.03	

Note. IR = immediate recall, DR = delayed recall, RT = recognition, CCI = central coherence index, VC1-4 = Verbal Fluency conditions 1-4, CW3-4 = Color-Word Interference conditions 3 and 4, TT = Tower of London test, TMT = Trails Making Task, BSAT = Brixton Spatial Anticipation Test. † = Brown-Forsythe ANOVA. Significant *p*-values ($p < 0.05$) indicated in bold. All values are z-scores.

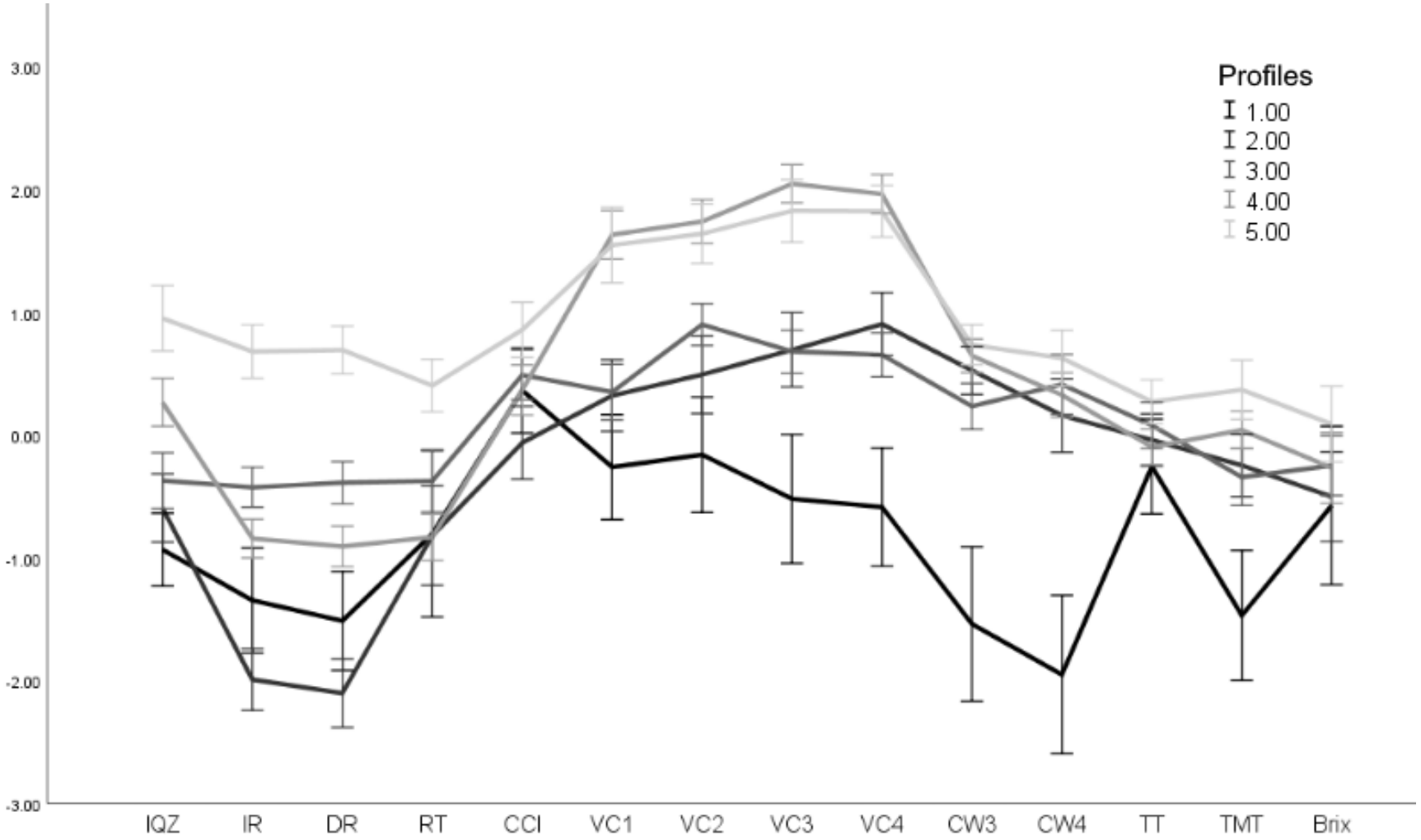


Figure 1. Standardised means between anorexia neuropsychological profiles. Error bars are 95% confidence intervals. IQZ = IQ, IR = immediate recall, DR = delayed recall, RT = recognition, CCI = Central Coherence Index, VC1-4 = Verbal Fluency Conditions 1-4, CW3-4 = Color-Word Interference conditions 3 and 4, TT = Tower Test, TMT = Trail Making Task.

Table 4.

Demographic and Clinical Variables Between Anorexia Nervosa Profiles

	Profile										Test statistic			
	1 (*n* = 18)		2 (*n* = 41)		3 (*n* = 69)		4 (*n* = 76)		5 (*n* = 49)					
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	*p*	η_p^2	Post hoc
Age	16.24	1.25	15.31	2.09	15.85	1.68	15.95	1.65	15.08	1.80	3.11†	**0.017**	0.05	1 > 5 (*p* = 0.037)
BMI Centile	4.04	5.50	5.38	8.24	11.93	17.54	5.55	9.32	7.27	10.10	3.93†	**0.005**	0.06	3 > 1 (*p* = 0.030)
Eating Restraint	4.00	1.67	2.67	1.84	3.31	2.02	3.42	1.94	2.79	1.73	2.22	0.068	0.04	
Eating Concern	3.35	1.64	2.54	1.87	3.00	1.81	2.94	1.73	2.61	1.70	0.96	0.433	0.02	
Weight Concern	3.92	1.82	3.17	1.97	3.56	1.97	3.58	1.96	3.13	2.09	0.80	0.527	0.02	
Shape Concern	4.28	1.69	3.59	1.86	3.88	1.89	3.97	1.99	3.48	2.01	0.80	0.526	0.02	
Global AN pathology	3.88	1.40	2.98	1.60	3.44	1.75	3.45	1.77	3.03	1.70	1.30	0.272	0.02	
Depression *z*	2.37	1.26	1.61	1.38	1.83	1.27	2.17	1.06	1.61	1.46	2.61†	**0.037**	0.04	
Anxiety *z*	1.52	1.11	1.06	.97	1.32	1.17	1.52	1.01	1.18	1.28	1.51	0.199	0.02	
Obsessionality *z*	2.19	1.37	1.63	1.64	2.01	1.25	1.99	1.34	1.63	1.56	1.17†	0.326	0.02	

Note. BMI = body mass index, † = Brown-Forsythe ANOVA. Significant *p*-values ($p < 0.05$) indicated in bold.

were high-average (0 < *z* < 1), except verbal fluency scores, which were high (*z* > 1). The greatest differences between the profiles were found on measures of IQ, memory (IR, DR, RT), and verbal fluency (VC1, VC2, VC3, VC4). There was scant discrimination between the profiles on central coherence (CCI), planning (TT), and inhibition (BSAT). Excluding the globally impaired group, inhibition (CW3 and CW4) and set-shifting (TMT) were also poor discriminators between the profiles.

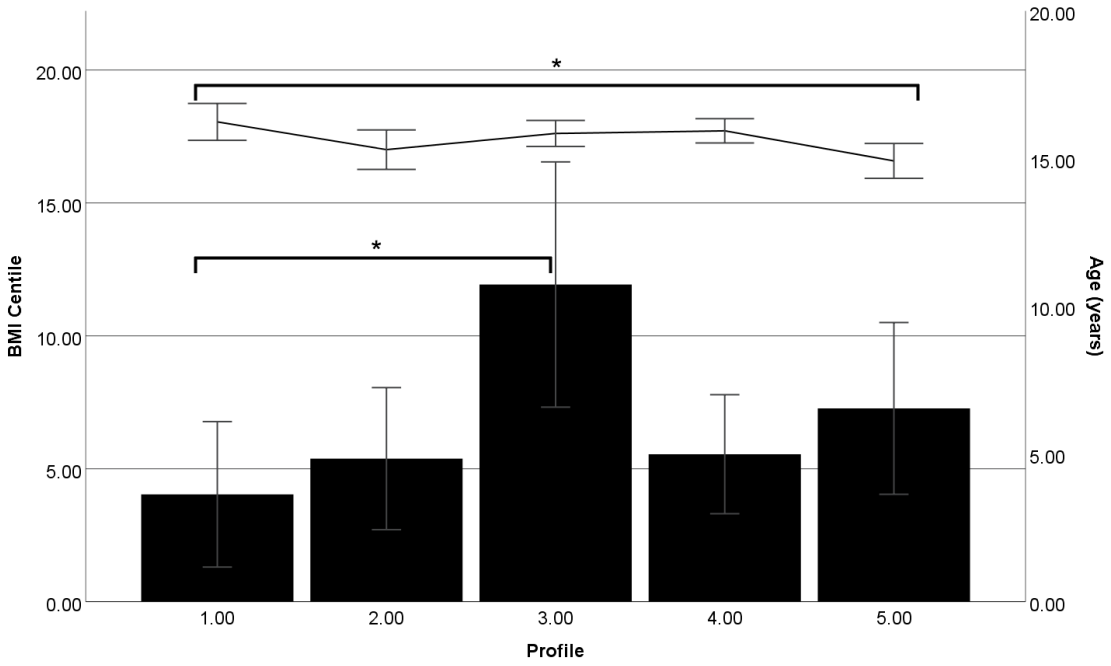


Figure 2. BMI centile (bars) and age (line) by anorexia nervosa neuropsychological profile. Error bars are 95% confidence intervals. * = significant (p < 0.05).

Multiple ANOVAs were performed to assess the differences between the five AN profiles on the clinical and demographic variables age, BMI, eating disorder pathology, anxiety, depression, and obsession-compulsion (Table 4). A one-way Brown-Forsythe ANOVA with age as the dependent variable and profile as the between-subjects factor was significant, $F(4, 191.737) = 3.107$, $p = 0.017$, $\eta_p^2 = 0.046$. Profile 5 (*high-average with strong* verbal) were younger

than profile 1 (*globally impaired*), with the Games-Howell post hoc test significant at $p = 0.037$ (Figure 2). A one-way Brown-Forsythe ANOVA with BMI centile as the dependent variable and profile as the between-subject factor was significant, $F(4, 163.139) = 3.925$, $p = 0.005$, $\eta_p^2 = 0.06$ (Figure 2). Profile 1 (*globally* impaired) had lower BMI centiles than profile 3 (*average performance*), with the Games-Howell test significant at $p = 0.03$ (Figure 2). A one-way Brown-Forsythe ANOVA with depression as the dependent variable and profile as the between-subjects factor was significant, $F(4, 161.786) = 2.730$, $p = 0.035$, $\eta_p^2 = 0.042$. Profile 4 (*visuo-verbal discrepancy*) were more depressed than profile 5 (*high-average with strong verbal*), though the Games-Howell post hoc test was non-significant ($p = 0.138$).

Healthy Control Profiles

Four profiles provided the optimal LPA solution to the HC neuropsychological test data (Table 5). A five-profile solution failed to replicate at the enhanced STARTS values (100-20) used for the AN LPA analysis. The posterior probabilities for assignment to each profile were very high (*p*s > 0.9), indicating a high degree of specificity between the different profiles (Table 6).

Table 5.

Healthy Controls Latent Profile Models

		Fit statistic		Profile membership distribution			
Profile	BIC	BLRT	Entropy	1	2	3	4
1	6505.68	–3180.94	NA	170			
2	6390.31	–3084.73*	0.820	99	71		
3	6315.11	–3008.62*	0.866	47	89	34	
4	6249.32	–2937.21*	0.894	49	44	51	26

Note. BIC = Bayesian information criterion, BLRT = bootstrapped likelihood ratio test, * = *p*-value <0.05 indicating the k-profile model is superior than the *k*–1-profile model.

Table 6.

Four-profile Healthy Control Model Membership Posteriors

Profile	*n*	%	1	2	3	4
1	49	29	**0.922**	0.026	0.052	0.000
2	44	26	0.008	**0.969**	0.020	0.002
3	51	30	0.032	0.027	**0.937**	0.004
4	26	15	0.000	0.009	0.036	**0.955**

Note. Bold values report the average posterior probability associated with the profile to which each case was assigned.

Large effect sizes were observed (η_p^2 > 0.14) in every case except for central coherence index, recognition, Color-Word Interference conditions 3 and 4, Tower Test, Trail Making Task, and BSAT test (0.02 < η_p^2 < 0.09), where it was small or medium (Table 7). A plot of the neuropsychological performance of each HC profile is presented in Figure 2.

The results of the between-profile comparisons were used to name the four HC profiles. Profile 1 (*n* = 49) comprised average performance (–0.5 < *M* < 0.5) on all tests except the immediate (*M* = –0.83) and delayed recall (*M* = –0.85) conditions of the Rey memory test, where performance was somewhat weaker. Accordingly, profile 1 was named *HC memory weakness*. While like the AN profile *memory impaired*, the memory weakness observed in HC profile 1 was less significant than in AN *impaired memory*, where scores were more than 2 SD below the global mean. Profile 2 (*n* = 44) scored in the upper-average range (0 < *M* < 1) on all tests and produced greater IQ and memory scores than any other profile. Profile 2 was named *above-average*. Profile 3 (*n* = 51) was equivalent to profile 2, except that performance on the measures of IQ (*M* = –0.43), immediate recall (*M* = –1.51), delayed recall (*M* = –1.34), and recognition (*M* = – 0.50) were weaknesses rather than strengths. Profile 3 was named *HC discrepant performance*. Finally, profile 4, named *HC strong verbal*, produced

average scores in most tests ($0.5 < M < 1$), except the four verbal fluency conditions, where performance was high to extremely high ($1 < M < 2.41$).

Multiple ANOVAs were performed to assess the differences between the four HC neuropsychological profiles on clinical and demographic variables (Table 8). All tests were non-significant, though the ANOVA for obsessive-compulsive symptoms approached significance ($p = 0.077$).

Neural Network Analysis

Fifty-six combinations of parameters were tested (Appendix L). Optimal model performance was obtained with batch size 5, epochs 500, and two hidden layers (Figure 3).

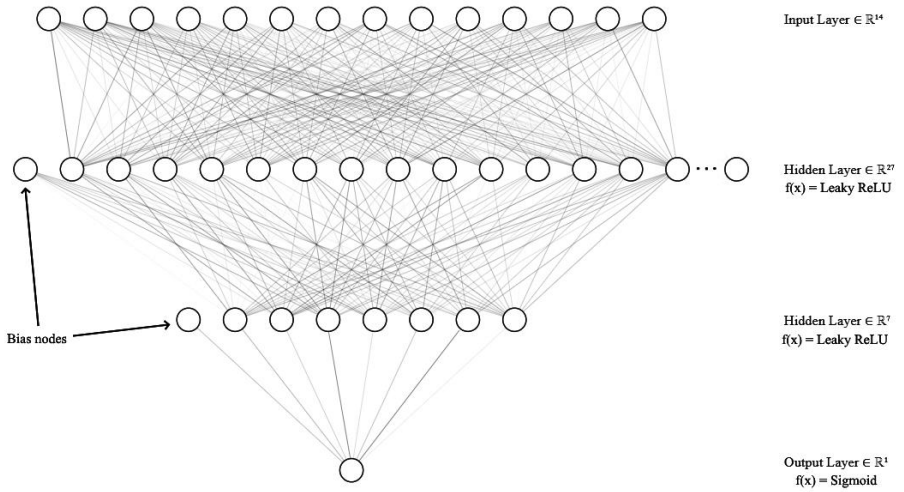


Figure 3. Final neural network model schematic. Not all ($n = 27$) neurons in the first hidden layer are represented. Bias nodes do not contribute to the total number of neurons in each layer. Greyscale lines represent different randomly initialised weight values.

Table 7.

Neuropsychological Performance Between Healthy Control Profiles

	Profile								Test statistic			
	1 (*n* = 49)		2 (*n* = 44)		3 (*n* = 51)		4 (*n* = 26)					
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	*p*	η_p^2	Post hoc
IQ	–0.04	0.77	0.65	0.99	–0.43	0.74	–0.20	0.66	14.66	**<0.001**	0.21	2 > 4, 3, 1
IR	–0.83	0.94	0.74	0.60	–1.51	0.82	–0.50	0.89	61.70	**<0.001**	0.53	2 > 4, 3, 1; 1 > 3
DR	–0.85	0.98	0.84	0.73	–1.34	0.78	–0.56	0.79	60.38	**<0.001**	0.52	2 > 4, 3, 1; 1, 4 > 3
RT	–0.41	1.10	0.33	1.00	–0.50	1.55	–0.09	0.77	4.63	**0.004**	0.08	2 > 1, 3
CCI	0.17	1.29	0.21	0.91	–0.42	0.78	–0.00	0.73	4.20	**0.007**	0.07	1, 2 > 3
VC1	–0.22	1.00	0.14	0.97	–0.16	0.81	0.91	0.55	11.82†	**<0.001**	0.16	4 > 1, 2, 3
VC2	0.14	0.77	0.78	0.94	0.95	0.82	2.35	0.73	40.74	**<0.001**	0.42	4 > 3, 2, 1; 3, 2 > 1
VC3	–0.27	0.65	0.97	0.49	0.85	0.49	2.41	0.51	141.86	**<0.001**	0.72	4 > 3, 2, 1; 3, 2 > 1
VC4	–0.14	0.58	1.08	0.42	1.07	0.38	2.18	0.53	142.83	**<0.001**	0.72	4 > 3, 2, 1; 3, 2 > 1
CW3	0.16	0.84	0.56	0.59	0.59	0.74	0.70	0.57	4.71	**0.003**	0.08	2, 3, 4 > 1
CW4	0.14	0.56	0.45	0.58	0.59	0.51	0.44	0.52	5.67	**0.001**	0.09	3 > 1
TT	0.32	0.92	0.24	0.84	0.03	0.60	0.17	0.71	1.25	0.293	0.02	
TMT	–0.14	0.82	0.24	0.68	0.41	0.68	–0.01	0.97	4.71	**0.003**	0.08	3 > 1
BSAT	0.09	0.83	0.06	0.74	–0.24	0.98	–0.13	0.84	1.59	0.193	0.03	

Note. IR = immediate recall, DR = delayed recall, RT = recognition, CCI = central coherence index, VC1-4 = Verbal Fluency conditions 1-4, CW3-4 = Color-Word Interference conditions 3 and 4, TT = Tower of London test, TMT = Trails Making Task, BSAT = Brixton Spatial Anticipation Test. † = Brown-Forsythe ANOVA. Significant *p*-values ($p < 0.05$) indicated in bold. All values are z-scores.

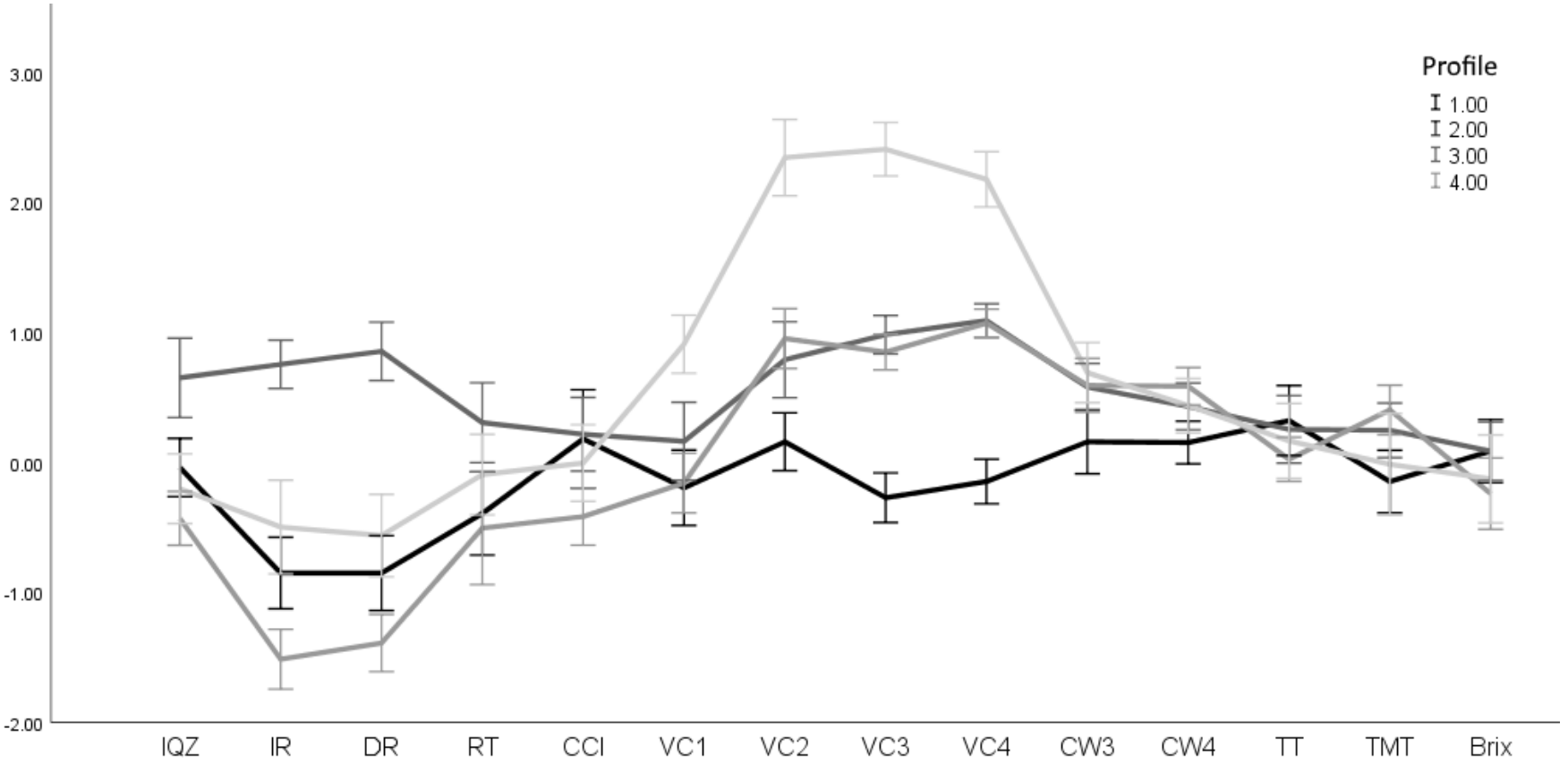


Figure 4. Standardised means between healthy control neuropsychological profiles. Error bars are 95% confidence intervals. IQZ = IQ, IR = immediate recall, DR = delayed recall, RT = recognition, CCI = Central Coherence Index, VC1-4 = Verbal Fluency conditions 1-4, CW3-4 = Color-Word Interference conditions 3 and 4, TT = Tower Test, TMT = Trail Making Task.

Table 8.

Demographic and Clinical Variables Between Healthy Control Profiles

	Profile								Test statistic			
	1 (*n* = 49)		2 (*n* = 44)		3 (*n* = 51)		4 (*n* = 26)					
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	*p*	η_p^2	Post hoc
Age	14.58	1.97	14.12	2.27	14.66	2.59	14.66	2.23	0.55	0.651	0.01	
BMI Centile	55.63	26.42	57.16	26.49	61.38	28.44	54.14	29.35	0.54	0.659	0.01	
Eating Restraint	0.49	0.73	0.44	0.76	0.71	1.02	0.55	1.05	0.86	0.563	0.02	
Eating Concern	0.43	0.67	0.22	0.56	0.31	0.53	0.49	0.90	1.30	0.276	0.02	
Weight Concern	0.91	1.10	0.66	1.00	0.91	1.05	1.24	1.62	1.24†	0.301	0.03	
Shape Concern	1.15	1.17	0.84	1.06	0.96	1.04	1.23	1.55	0.86	0.466	0.02	
EDE Global	0.74	0.85	0.56	0.79	0.71	0.80	0.91	1.20	0.87	0.457	0.02	
Depression *z*	–0.52	0.88	–0.75	0.66	–0.44	0.87	–0.32	0.88	1.80	0.149	0.03	
Anxiety *z*	–0.34	0.77	–0.70	0.72	–0.53	0.61	–0.62	0.75	2.14	0.097	0.04	
Obsessionality *z*	0.07	1.07	–0.28	0.89	–0.03	0.90	0.32	0.97	2.36	0.077	0.04	

Note. BMI = body mass index, EDE = Eating Disorder Examination, † = Brown-Forsythe ANOVA. Significant *p*-values ($p < 0.05$) indicated in bold.

The performance of the linear and non-linear models was not significantly different, $\tilde{t}(1, 5) = 0.150$, $p = 0.443$. Linear and non-linear model performance was compared when supplied with additional input variables: age; depression, anxiety, and obsessionality; eating disorder pathology; and BMI, to test whether these additional variables introduced non-linear structure. Only when every additional input variable was inputted did the non-linear model outperform the linear model (Table 9), demonstrating that neuropsychological performance, psychological variables, and eating disorder pathology are not better described by a non-linear model.

Table 9.

Linear (Identity) vs. Non-Linear (ReLU) Model Performance

		Linear		Non-linear		Test statistic	
M	Data	*M*	*SD*	*M*	*SD*	*t*	*p*
1	Neuro only	71.07	2.73	71.12	3.59	0.15	0.443
2	1 + age	70.71	1.43	70.00	3.18	0.89	0.208
3	2 + anxiety	84.56	1.78	84.98	2.49	1.93	0.056
4	3 + depression	86.86	2.08	85.21	1.24	1.82	0.065
5	4 + obsessionality	86.04	2.97	85.86	1.88	1.71	0.074
6	5 + EDE	86.80	2.66	86.80	1.91	1.58	0.087
7	6 + BMI	90.36	1.72	92.72	1.35	2.46	**0.029**

Note. ReLU = rectifier linear unit, scores are cases (%) correctly predicted, neuro = IQ, immediate recall, delayed recall, recognition, verbal fluency, central coherence, and set-shifting; EDE = eating disorder examination global and subscales; BMI = body mass index.

Discussion

The present study sought to identify latent profiles of neuropsychological function within a large sample of AN patients. It was hoped the profiles could be linked to clinical and demographic variables that would assist our understanding

and treatment of the disorder. A NN was also used to ascertain whether previous failures to link neuropsychological performance with clinical and demographic variables was due to the relationship between them being non-linear.

LPA Profiles

The first research question concerned latent profiles within the neuropsychological dataset. Latent profile analysis revealed that patients with AN could be grouped into five profiles determined by neuropsychological performance. The five profiles were well populated and substantially differed with one another on most of the neuropsychological variables. The tests that contributed least to discriminating the five profiles were the Tower Test of planning, the Brixton Spatial Anticipation Test of flexibility, and the Rey Complex Figure Test of central coherence. In each case, the variance between the five profiles was less in these three functions than the others and unrelated to ability in functions that might be assumed to draw upon similar abilities. For example, to succeed at the Tower Test participants need to visualise and remember a sequence of steps, looking ahead to avoid error. Participants that performed well on the measures of visual memory might be expected to perform well in the Tower Test, but the data do not bear this out. The correlations between the visual memory measures and Tower Test were significant, although with very weak effect sizes (Appendix M). The variability in performance clearly visible among the five AN profiles is masked when AN patients are treated as a homogeneous group. It was surprising that central coherence contributed so little to discriminating the profiles. It may be that central coherence does so little because performance was relatively uniform among AN patients, suggesting central coherence is a core neuropsychological construct within AN,

although data from this study are at odds with the literature in indicating that AN patients had superior central coherence than HCs (Appendix N).

When the *globally impaired* profile 1 were excluded, Color-Word Interference measures of inhibition and the Trail Making Task measure of set-shifting were also poor discriminators between the profiles. The primary and strongest differences between AN patients were between IQ, visual memory, and verbal fluency. The *globally impaired* profile 1 was the smallest latent profile among the AN patients and may be artificially shrunk by the recruitment used in the study, which excluded participants with an IQ estimate below 85 (Rose et al., 2016).

The comparisons between latent profiles provided support to research question two that neuropsychologically-based AN profiles would differ on clinical meaningful variables. The five profiles were equivalent on measures of weight concern, shape concern, eating concern, or global eating disorder pathology. Only eating restraint approached significance, with the *globally impaired* profile 1 showing greater restraint than the *memory impaired* profile 2. Of the psychological variables, obsessionality and anxiety were comparable across the five profiles. The failure to detect differences between the profiles is unlikely to be due to sampling effects. Eating disorder pathology among the AN sample was high (Appendix O) and much greater than community norms (Carter, Stewart, & Fairburn, 2001). A significant difference emerged between the five profiles and depression, though the post hoc Bonferroni comparison was non-significant. G*lobally impaired* profile 1 was older than the *high average with strong verbal* profile 5, with a near-moderate effect size. BMI centile also differed, with *globally impaired* profile 1 recording lower BMI centiles than the *average performance* profile 3. That both significant

differences involved profile 1 lends weight to the findings. The observed differences between clinical and demographic variables of the five neuropsychological profiles provides some evidence to support research question two. It is possible that a larger sample size may produce significant differences in depression, although the feasibility of doing so is questionable given the number of participants required.

There were some similarities and differences between the clusters identified by Rose et al. (2016) and those identified within the same data in the present study. Rose et al. described three clusters, a low average to average cluster (19%), a discrepant visual-verbal cluster (33%), and a strong verbal cluster (48%). The discrepant visual-verbal cluster was like *visual-verbal discrepancy* profile 4. Both profiles contained participants with average scores except on verbal tests, at which they performed better than average, and visuospatial memory tests, where performance was somewhat worse than average. While the *globally impaired* profile 1 described by here appears like the low average to average cluster reported by Rose et al., the magnitude of the neuropsychological weaknesses was much greater in the *globally impaired* profile. The strong verbal cluster identified by Rose et al. was not equivalent to any of the profiles that emerged from the latent profile analysis. It was most like the *high-average with strong verbal* profile 5, though non-verbal scores were not as high, with many close to the mean. Unlike the present research, the clusters identified by Rose et al. were not associated with any clinical or demographic variable. This suggests that there may be neuropsychological profiles within AN patients that do associate with clinical or

demographic variables; however, the number and strength of those relationships was low and weak.

Based on a much smaller range of neuropsychological tests – Wisconsin Card Sorting Task (problem solving, perseveration), Brixton Spatial Anticipation Test (set-shifting, cognitive flexibility), Rey Complex Figure Test (central coherence), and the Reading the Mind in the Film Task (emotional theory of mind) – and using cluster analysis, Renwick et al. (2015) also discovered three clusters of neuropsychological performance. One closely resembled the *globally impaired* profile 1, with all scores impaired ($-2.41 < z < -1.07$) except the Brixton test ($z = -0.11$). Neither other group was comparable with the profiles identified in the present research. Cluster 2 showed impaired set-shifting and central coherence with average perseveration and emotional theory of mind. Cluster 1 showed impaired central coherence and average emotional theory of mind, set-shifting, and perseveration. As with Rose et al. (2015), Renwick et al. were unable to identify any clinical or demographic differences between their three clusters. The only difference that was close to achieving significance was IQ ($p = 0.06$), for which cluster 1 had a higher score than cluster 3 ($d = 0.7$).

NN Non-linearity

The third research question, whether the neuropsychological data contains non-linear structure, has been answered in the negative. While a non-linear NN was capable of accurately predicting 71% of cases under the most stringent conditions (i.e. predicting diagnostic status using only neuropsychological test data), it was not superior to a simple linear model. Adding dimensions to the dataset did not alter the linear nature of the relationship until the final variable, BMI,

was entered, at which point the non-linear model became significantly superior to the linear model. Unfortunately, it is impossible to assess the contributions of input variables in a NN, so it is difficult to interpret this finding. Alone, BMI is likely to possess a non-linear, sigmoidal, relationship with diagnostic status as the diagnosis of AN is dependent on a threshold BMI around 17.5 kg/m^2, although the diagnosis is slightly different in children and young people. Whether BMI interacts with other input variables in the model is unknown.

Pulling together these findings, it appears possible to reliably identify a *globally impaired group* among AN patients, who show deteriorated performance in most neuropsychological tests. However, it is important to note that this performance impairment is only relative to other patients with AN. Objectively, the *globally impaired* profile 1 identified in the present research still had an average IQ (M = 96.3, SD = 7.07). As Renwick et al. did not assess verbal or visuospatial performance, it is impossible to draw any further links between their clusters and those identified by Rose et al. or the present research. The *globally impaired* profile 1 was older than the *high average with strong verbal* profile 5, which may mean their disorder developed later or remained undetected and untreated for longer, thereby becoming more severe. Profile 1 also had lower BMI centiles than the *average performance* profile 3, suggesting impaired neuropsychological presentation may be linked to greater pathology. Profile 1 did score more highly than the other groups on the measure of eating restraint, though the result was marginally non-significant (p = 0.068). The samples were generally recruited from inpatient treatment facilities (Rose et al., 2016). It may be that if recruitment

included more community patients with less severe eating disorder pathology, that the differences between them and profile 1 may be more significant.

The noradrenergic model (Nunn et al., 2012) predicts that anxiety is a key psychological mediator of self-starvation behaviour in some cases of AN. The model claims that anxiety is produced when the internal representations of the body produced by interoceptive and proprioceptive awareness desynchronise. One purpose of looking for neuropsychological profiles within AN is to test whether any of the profiles differ in the amount of anxiety reported by patients. The results of the present research suggest that is not the case, with every neuropsychological profile characterised by equivalent levels of anxiety. However, imprecision may explain the findings. Anxiety is treated as a bio-cognitive construct. Traditionally, anxiety was conceptualised as a biological process: the body's hypothalamic-pituitary-adrenal (HPA) axis response to threat. Modern parlance has repurposed the word stress to describe the state of HPA activation and HPA activators (i.e. stressors). Assessing the level of HPA activation in AN patients (e.g. querying striated and smooth muscle tension, cognitive-perceptual disturbance, and conversion) or with a biological marker such as salivary cortisol (Törnhage, 2009) may be a better measure of the anxiety that Nunn et al. predict underlies self-starvation behaviour.

Clinical Implications

Implications for assessment. It is costly and time consuming to administer a full neuropsychological battery to AN patients. If clinicians wish to assess AN patients' neuropsychological performance for treatment purposes but have limited time to do so, the results of the present research suggest they should use the

Delayed Recall Trial (DR) of the Rey Complex Figure Test and Total Switching Accuracy (VC4) from the DKEFS Verbal Fluency test. These two measures – one a visual task, the other a verbal one – differ with the greatest effect sizes among the five profiles detected by the latent profile analysis conducted in the present research and so can provide a quick indicator of the patient's verbal and visual abilities.

Implications for treatment. New research that aims to translate the findings from AN neuropsychological research into practice is the use of cognitive remediation therapy (CRT), an approach designed to promote cognitive flexibility. The findings that AN patients often have deficits in central coherence and set-shifting are thought to hinder treatment for the disorder (Tchanturia et al., 2013). Early research suggested patients who received CRT improve on measures of set-shifting and central coherence (Abbate-Daga, Buzzichelli, Marzola, Amianto, & Fassino, 2012; Tchanturia et al., 2008) and that improvements may be associated with clinically relevant variables such as impulse regulation and interoceptive awareness (Abbate-Daga et al., 2012). Research investigating the neuropsychological profiles of AN has raised the prospect that for CRT to be effective, it should target person-specific neuropsychological impairments. The findings of the present research suggest this may not be necessary. The neuropsychological profiles within AN patients were relatively homogeneous in terms of set-shifting and central coherence, with effect size differences between them being small. Only the *globally impaired* profile 5 differed markedly on these measures. The present research suggests it may be empirically justifiable to adopt a manualised, as opposed to personalised, approach to CRT interventions.

Patients first admitted for AN treatment that are older (15 years and older) and lighter (BMI centile 9 or lower) and may have, relative to other AN patients, an impaired cognitive profile. These impairments may be greater in the clinic compared to the research lab as participants with a prorated IQ below 85 were removed from the present research. Rather than the result of an impaired cognitive profile, an alternative explanation for the differences in age and BMI may be that the *globally impaired* profile 5 patients may have starved for longer before presenting to services and therefore may have suffered some or greater brain damage than patients from the four other neuropsychological profiles. The outcomes of treatment may be affected by the poorer cognitive performance of this group. A longer duration or more intensive treatment may be indicated for these patients, as may more robust monitoring post discharge to spot early signs of relapse.

Implications for services. There have been calls for services to invest in specialist neuropsychological assessment and neuropsychological interventions such as CRT. The present study suggests this may be premature, especially given the paucity of qualified neuropsychologists required to perform such assessment and intervention. Services may benefit more from investing in treatment approaches that are informed by contemporary neuropsychological understanding of AN. The Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA; Wade, Treasure, & Schmidt, 2011) is a good example of such an intervention. It includes modules that cover thinking styles and cognitive flexibility in a format that can be delivered by a wide range of mental health professionals. MANTRA was

recently recognised as a recommended psychological intervention for adults with AN by the National Institute for Health and Care Excellence (2017).

Implications for research. Evidence of neuropsychological dysfunction in AN has prompted two interpretations: that neuropsychological dysfunction either predates starvation behaviour or is its result (Holliday, Tchanturia, Landau, Collier, & Treasure, 2005; Roberts, Tchanturia, & Treasure, 2013). A third interpretation needs consideration. Impaired neuropsychological test performance may merely describe the psychological features of AN (cognitive and behavioural dysfunction) in neuropsychological terms. While enticing, such descriptions may not offer any additional explanatory power. If AN is characterised by the obsessive pursuit of a goal (a behavioural description), then it should not surprise us that when tested neuropsychologically their results indicate weak central coherence. Weak central coherence may literally be part of the neuropsychological description of obsessionality. Although a very difficult theory to test, an implication of it would be that there is little utility in identifying the causal relationship between neuropsychological functioning and the features of AN and that research effort current expended trying to establish such causality could be better discharged elsewhere.

The present research used NNs to look for non-linearity in the dataset. While a cutting-edge technology, the output of the NN from the large dataset reported here was disappointing. NNs thrive when presented with arbitrarily complex stimuli that cannot be easily decomposed, for example an image or audio clip, each containing millions of data points. Applying the method to psychological datasets as was attempted here, appears unlikely to provide additional explanatory

power that could not be modelled *apriori* with conventional non-hypothesis significance testing and should, therefore, be avoided.

Conclusion

The present study attempted to identify clinically meaningful neuropsychologically-defined subgroups in AN using objective statistical and non-linear modelling approaches. The findings indicate five neuropsychological profiles were present. One profile, *globally impaired,* had a lower BMI centile than the *average performance* profile and was older than the *high-average with strong verbal* profile. The NN analysis revealed that non-linear modelling of AN neuropsychological ability was unhelpful. These findings provide tentative support for the clinical utility of neuropsychologically-derived subgroups of AN. That anxiety did not differ between profiles challenges a prediction of the theoretically-derived noradrenergic model of AN, inviting further refinement or abandonment of brain-based theories in pursuit of a person-based understanding for this complex and poorly-understood disorder.

References

Abbate-Daga, G., Buzzichelli, S., Marzola, E., Amianto, F., & Fassino, S. (2012). Effectiveness of cognitive remediation therapy (CRT) in anorexia nervosa: a case series. *Journal of Clinical and Experimental Neuropsychology*, *34*(10), 1009–1015. https://doi.org/10.1080/13803395.2012.704900

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington: Author.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.

Bentz, M., Jepsen, J. R. M., Kjaersdam Telléus, G., Moslet, U., Pedersen, T., Bulik, C. M., & Plessen, K. J. (2017). Neurocognitive functions and social functioning in young females with recent-onset anorexia nervosa and recovered individuals. *Journal of Eating Disorders*, *5*, 5. https://doi.org/10.1186/s40337-017-0137-3

Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An Introduction to Latent Variable Mixture Modeling (Part 1): Overview and Cross-Sectional Latent Class and Latent Profile Analyses. *Journal of Pediatric Psychology*, *39*(2), 174–187. https://doi.org/10.1093/jpepsy/jst084

Bishop, C. M. (2006). *Pattern recognition and machine* learning. New York: Springer.

Booth, R. D. L. (2006). *Local-global processing and cognitive style in autism spectrum disorders and typical development* (Ph.D., University of London).

Retrieved from https://kclpure.kcl.ac.uk/portal/en/theses/localglobal-processing-and-cognitive-style-in-autism-spectrum-disorders-and-typical-development(99cff8ab-25d8-4530-8e66-dc089e8db9fd).html

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests.* Bury St Edmunds: Thames Valley Test Company.

Brewerton, T. D., Frampton, I., & Lask, B. (2009). The neurobiology of anorexia nervosa. *US Psychiatry*, *2*(1), 57–60.

Bryant-Waugh, R. (2000). Overview of the eating disorders. In Lask, B., & Bryant-Waugh, R. (Eds.) *Anorexia nervosa and related eating disorders in childhood and adolescence* (pp. 27–40). Hove: Psychology Press.

Carter, J. C., Stewart, D. A., & Fairburn, C. G. (2001). Eating disorder examination questionnaire: norms for young adolescent girls. *Behaviour Research and Therapy*, *39*(5), 625–632. https://doi.org/10.1016/S0005-7967(00)00033-4

Castro-Fornieles, J., Bargalló, N., Lázaro, L., Andrés, S., Falcon, C., Plana, M. T., & Junqué, C. (2009). A cross-sectional and follow-up voxel-based morphometric MRI study in adolescent anorexia nervosa. *Journal of Psychiatric Research*, *43*(3), 331–340. https://doi.org/10.1016/j.jpsychires.2008.03.013

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Cole, T. J., Freeman, J. V., & Preece, M. A. (1995). Body mass index reference curves for the UK, 1990. *Archives of Disease in Childhood*, *73*(1), 25–29.

Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, *6*(1), 1–8. https://doi.org/10.1002/1098-108X(198701)6:1<1::AID-EAT2260060102>3.0.CO;2-9

Danner, U. N., Sanders, N., Smeets, P. A. M., van Meer, F., Adan, R. A. H., Hoek, H. W., & van Elburg, A. A. (2012). Neuropsychological weaknesses in anorexia nervosa: set-shifting, central coherence, and decision making in currently ill and recovered women. *The International Journal of Eating Disorders*, *45*(5), 685–694. https://doi.org/10.1002/eat.22007

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System.* London: Pearson Assessment.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, *10*(7), 1895–1923. https://doi.org/10.1162/089976698300017197

Duchesne, M., Mattos, P., Fontenelle, L. F., Veiga, H., Rizo, L., & Appolinario, J. C. (2004). [Neuropsychology of eating disorders: a systematic review of the literature]. *Revista Brasileira De Psiquiatria*, *26*(2), 107–117. https://doi.org//S1516-44462004000200008

Dziak, J. J., Lanza, S. T., & Tan, X. (2014). Effect Size, Statistical Power and Sample Size Requirements for the Bootstrap Likelihood Ratio Test in Latent

Class Analysis. *Structural Equation Modeling : A Multidisciplinary Journal*, *21*(4), 534–552. https://doi.org/10.1080/10705511.2014.919819

Fowler, L., Blackwell, A., Jaffa, A., Palmer, R., Robbins, T. W., Sahakian, B. J., & Dowson, J. H. (2006). Profile of neurocognitive impairments associated with female in-patients with anorexia nervosa. *Psychological Medicine*, *36*(4), 517–527. https://doi.org/10.1017/S0033291705006379

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., … Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry*, *46*(11), 1006–1011. https://doi.org/10.1001/archpsyc.1989.01810110048007

Green, M. W., Elliman, N. A., Wakeling, A., & Rogers, P. J. (1996). Cognitive functioning, weight change and therapy in anorexia nervosa. *Journal of Psychiatric Research*, *30*(5), 401–410. https://doi.org/10.1016/0022-3956(96)00026-X

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Harrison, A., Tchanturia, K., Naumann, U., & Treasure, J. (2012). Social emotional functioning and cognitive styles in eating disorders. *The British Journal of Clinical Psychology*, *51*(3), 261–279. https://doi.org/10.1111/j.2044-8260.2011.02026.x

Harrison, A., Tchanturia, K., & Treasure, J. (2011). Measuring state trait properties of detail processing and global integration ability in eating disorders. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry*, *12*(6), 462–472. https://doi.org/10.3109/15622975.2010.551666

Holliday, J., Tchanturia, K., Landau, S., Collier, D., & Treasure, J. (2005). Is impaired set-shifting an endophenotype of anorexia nervosa? *The American Journal of Psychiatry*, *162*(12), 2269–2275. https://doi.org/10.1176/appi.ajp.162.12.2269

Jones, B. P., Duncan, C. C., Brouwers, P., & Mirsky, A. F. (1991). Cognition in eating disorders. *Journal of Clinical and Experimental Neuropsychology*, *13*(5), 711–728. https://doi.org/10.1080/01688639108401085

Kemps, E., Tiggemann, M., Wade, T., Ben-Tovim, D., & Breyer, R. (2006). Selective working memory deficits in anorexia nervosa. *European Eating Disorders Review*, *14*(2), 97–103. https://doi.org/10.1002/erv.685

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal*, *17*(6), 441–458.

Key, A., O'Brien, A., Gordon, I., Christie, D., & Lask, B. (2006). Assessment of neurobiology in adults with anorexia nervosa. *European Eating Disorders Review*, *14*(5), 308–314. https://doi.org/10.1002/erv.696

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *International Conference on Learning Representations 2015*. Retrieved from https://arxiv.org/pdf/1412.6980.pdf

Kim, Y.-R., Lim, S.-J., & Treasure, J. (2011). Different patterns of emotional eating and visuospatial deficits whereas shared risk factors related with social support between anorexia nervosa and bulimia nervosa. *Psychiatry Investigation*, *8*(1), 9–14. https://doi.org/10.4306/pi.2011.8.1.9

Kovacs, M. (1992). *Children's Depression Inventory*. North Tonawanda, NY: Multi-Health Systems, Inc.

Lang, K., Lopez, C., Stahl, D., Tchanturia, K., & Treasure, J. (2014). Central coherence in eating disorders: An updated systematic review and meta-analysis. *The World Journal of Biological Psychiatry*, *15*(8), 586–598. https://doi.org/10.3109/15622975.2014.909606

Lang, K., & Tchanturia, K. (2014). A systematic review of central coherence in young people with anorexia nervosa. *Journal of Child and Adolescent Behavior*, 1–5. https://doi.org/10.4172/2375-4494.1000140

Lao-Kaim, N. P., Giampietro, V. P., Williams, S. C. R., Simmons, A., & Tchanturia, K. (2014). Functional MRI investigation of verbal working memory in adults with anorexia nervosa. *European Psychiatry*, *29*(4), 211–218. https://doi.org/10.1016/j.eurpsy.2013.05.003

Lena, S. M., Fiocco, A. J., & Leyenaar, J. K. (2004). The role of cognitive deficits in the development of eating disorders. *Neuropsychology Review*, *14*(2), 99–113. https://doi.org/10.1023/B:NERV.0000028081.40907.de

Lopez, C., Stahl, D., & Tchanturia, K. (2010). Estimated intelligence quotient in anorexia nervosa: a systematic review and meta-analysis of the literature. *Annals of General Psychiatry*, *9*, 40. https://doi.org/10.1186/1744-859X-9-40

Lopez, C., Tchanturia, K., Stahl, D., Booth, R., Holliday, J., & Treasure, J. (2008). An examination of the concept of central coherence in women with anorexia nervosa. *The International Journal of Eating Disorders*, *41*(2), 143–152. https://doi.org/10.1002/eat.20478

Lopez, C., Tchanturia, K., Stahl, D., & Treasure, J. (2009). Weak central coherence in eating disorders: a step towards looking for an endophenotype of eating disorders. *Journal of Clinical and Experimental Neuropsychology*, *31*(1), 117–125. https://doi.org/10.1080/13803390802036092

Maas, A., Hannun, A. & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. *International Conference on Machine Learning 2013*. Retrieved from https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf

McLachlan, G. & Peel, D. (2000). *Finite mixture models.* New York, NY: John Wiley.

Mesulam, M., & Mufson, E. (1988). The insula of Reil in man and monkey. In E. G. Jones, & A. Peters (Eds.), *Cerebral Cortex* (pp. 219–220). New York: Associationand Auditory Cortex Plenium Press.

Meyers, J. E. & Meyers, K. R. (1995). *Rey Complex Figure Test and recognition trial: Professional manual*. Odessa: Psychological Assessment Resources.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide: Statistical analysis with latent variables (7th ed.)*. Los Angeles, CA: Authors.

National Institute for Health and Care Excellence. (2017). *Eating disorders: Recognition and treatment*. Retrieved from https://www.nice.org.uk/guidance/ng69

Nunn, K., Frampton, I., Gordon, I., & Lask, B. (2008). The fault is not in her parents but in her insula--a neurobiological hypothesis of anorexia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *16*(5), 355–360. https://doi.org/10.1002/erv.890

Nunn, K., Frampton, I., & Lask, B. (2012). Anorexia nervosa--a noradrenergic dysregulation hypothesis. *Medical Hypotheses*, *78*(5), 580–584. https://doi.org/10.1016/j.mehy.2012.01.033

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology*, *32*(1), 8–47. https://doi.org/10.1016/j.cedpsych.2006.10.003

Rauch, S. L., & Savage, C. R. (1997). Neuroimaging and neuropsychology of the striatum: Bridging basic science and clinical practice. *The Psychiatric Clinics of North America*, *20*(4), 741–768.

Renwick, B., Musiat, P., Lose, A., DeJong, H., Broadbent, H., Kenyon, M., … Schmidt, U. (2015). Neuro- and social-cognitive clustering highlights distinct profiles in adults with anorexia nervosa. *The International Journal of Eating Disorders*, *48*(1), 26–34. https://doi.org/10.1002/eat.22366

Roberts, M. E., Tchanturia, K., Stahl, D., Southgate, L., & Treasure, J. (2007). A systematic review and meta-analysis of set-shifting ability in eating disorders. *Psychological Medicine*, *37*(8), 1075–1084. https://doi.org/10.1017/S0033291707009877

Roberts, M. E., Tchanturia, K., & Treasure, J. L. (2013). Is attention to detail a similarly strong candidate endophenotype for anorexia nervosa and bulimia nervosa? *The World Journal of Biological Psychiatry*, *14*(6), 452–463. https://doi.org/10.3109/15622975.2011.639804

Rose, M., Davis, J., Frampton, I., & Lask, B. (2011). The Ravello Profile: development of a global standard neuropsychological assessment for young people with anorexia nervosa. *Clinical Child Psychology and Psychiatry*, *16*(2), 195–202. https://doi.org/10.1177/1359104511401191

Rose, M., Frampton, I., & Lask, B. (2012). A case series investigating distinct neuropsychological profiles in children and adolescents with anorexia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *20*(1), 32–38. https://doi.org/10.1002/erv.1095

Rose, M., Stedal, K., Reville, M. C., van Noort, B. M., Kappel, V., Frampton, I., … Lask, B. (2016). Similarities and differences of neuropsychological profiles in children and adolescents with anorexia nervosa and healthy controls using

cluster and discriminant function analyses. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*. https://doi.org/10.1093/arclin/acw068

Rubinstein, R. Y., & Kroese, D. P. (2004). *The cross-entropy method: A unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning.* New York: Springer-Verlag.

Schmidt, U., & Treasure, J. (2006). Anorexia nervosa: valued and visible. A cognitive-interpersonal maintenance model and its implications for research and practice. *The British Journal of Clinical Psychology*, *45*(Pt 3), 343–366.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.2307/2958889

Shafran, R., & Somers, J. (1998). Treating adolescent obsessive-compulsive disorder: applications of the cognitive theory. *Behaviour Research and Therapy*, *36*(1), 93–97.

Shafran, R., Frampton, I., Heyman, I., Reynolds, M., Teachman, B., & Rachman, S. (2003). The preliminary development of a new self-report measure for OCD in young people. *Journal of Adolescence*, *26*(1), 137–142. https://doi.org/10.1016/S0140-1971(02)00083-0

Shelley, B. P., & Trimble, M. R. (2004). The insular lobe of Reil--its anatamico-functional, behavioural and neuropsychiatric attributes in humans--a review. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry*, *5*(4), 176–200.

Sherman, B. J., Savage, C. R., Eddy, K. T., Blais, M. A., Deckersbach, T., Jackson, S. C., … Herzog, D. B. (2006). Strategic memory in adults with anorexia nervosa: are there similarities to obsessive compulsive spectrum disorders? *The International Journal of Eating Disorders*, *39*(6), 468–476. https://doi.org/10.1002/eat.20300

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*(6), 553–566. https://doi.org/10.1080/09638230500347541

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R. & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto: Consulting Psychologists Press.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

Stanley, L., Kellermanns, F. W., & Zellweger, T. M. (2017). Latent profile analysis: Understanding family firm profiles. *Family Business Review*, *30*(1), 84–102. https://doi.org/10.1177/0894486516677426

Stedal, K., Frampton, I., Landrø, N. I., & Lask, B. (2012). An examination of the Ravello profile — a neuropsychological test battery for anorexia nervosa. *European Eating Disorders Review*, *20*(3), 175–181. https://doi.org/10.1002/erv.1160

Stedal, K., Rose, M., Frampton, I., Landrø, N. I., & Lask, B. (2012). The neuropsychological profile of children, adolescents, and young adults with anorexia nervosa. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *27*(3), 329–337. https://doi.org/10.1093/arclin/acs032

Steinglass, J. E., Walsh, B. T., & Stern, Y. (2006). Set shifting deficit in anorexia nervosa. *Journal of the International Neuropsychological Society: JINS*, *12*(3), 431–435.

Steinglass, J., & Walsh, B. T. (2006). Habit learning and anorexia nervosa: a cognitive neuroscience hypothesis. *The International Journal of Eating Disorders*, *39*(4), 267–275. https://doi.org/10.1002/eat.20244

Szmukler, G. I., Andrewes, D., Kingston, K., Chen, L., Stargatt, R., & Stanley, R. (1992). Neuropsychological impairment in anorexia nervosa: Before and after refeeding. *Journal of Clinical and Experimental Neuropsychology*, *14*(2), 347–352. https://doi.org/10.1080/01688639208402834

Tchanturia, K., Anderluh, M. B., Morris, R. G., Rabe-Hesketh, S., Collier, D. A., Sanchez, P., & Treasure, J. L. (2004). Cognitive flexibility in anorexia nervosa and bulimia nervosa. *Journal of the International Neuropsychological Society: JINS*, *10*(4), 513–520. https://doi.org/10.1017/S1355617704104086

Tchanturia, K., Campbell, I. C., Morris, R., & Treasure, J. (2005). Neuropsychological studies in anorexia nervosa. *International Journal of Eating Disorders*, *37*(S1), S72–S76. https://doi.org/10.1002/eat.20119

Tchanturia, K., Davies, H., Lopez, C., Schmidt, U., Treasure, J., & Wykes, T. (2008). Neuropsychological task performance before and after cognitive remediation in anorexia nervosa: A pilot case-series. *Psychological Medicine*, *38*(9), 1371–1373. https://doi.org/10.1017/S0033291708003796

Tchanturia, K., Lloyd, S., & Lang, K. (2013). Cognitive remediation therapy for anorexia nervosa: Current evidence and future research directions. *International Journal of Eating Disorders*, *46*(5), 492–495. https://doi.org/10.1002/eat.22106

Telléus, G. K., Jepsen, J. R., Bentz, M., Christiansen, E., Jensen, S. O. W., Fagerlund, B., & Thomsen, P. H. (2015). Cognitive Profile of Children and Adolescents with Anorexia Nervosa. *European Eating Disorders Review*, *23*(1), 34–42. https://doi.org/10.1002/erv.2337

Törnhage, C. J. (2009). Salivary Cortisol for Assessment of Hypothalamic-Pituitary-Adrenal Axis Function. *Neuroimmunomodulation*, *16*(5), 284–289. https://doi.org/10.1159/000216186

Touyz, S. W., Beumont, P. J. V., & Johnstone, L. C. (1986). Neuropsychological correlates of dieting disorders. *International Journal of Eating Disorders*, *5*(6), 1025–1034. https://doi.org/10.1002/1098-108X(198609)5:6<1025::AID-EAT2260050606>3.0.CO;2-T

Wade, T., Treasure, J., & Schmidt, U. (2011). A case series evaluation of the Maudsley Model for treatment of adults with anorexia nervosa. *European Eating Disorders Review, 19*(5), 382-389. https://doi.org/10.1002/erv.1078

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence.* San Antonio, TX: The Psychological Corporation.

Weider, S., Indredavik, M. S., Lydersen, S., & Hestad, K. (2014). Intellectual function in patients with anorexia nervosa and bulimia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *22*(1), 15–24. https://doi.org/10.1002/erv.2260

Appendices

Appendix A. Eating Disorder Examination Questionnaire

EATING QUESTIONNAIRE

Instructions: The following questions are concerned with the past four weeks (28 days) only. Please read each question carefully. Please answer all of the questions. Please only choose one answer for each question. Thank you.

Questions 1 to 12: Please circle the appropriate number on the right. Remember that the questions only refer to the past four weeks (28 days) only.

	On how many of the past 28 days ……	days	days	days	days	days	days	day
1	Have you been deliberately trying to limit the amount of food you eat to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
2	Have you gone for long periods of time (8 waking hours or more) without eating anything at all in order to influence your shape or weight?	0	1	2	3	4	5	6
3	Have you tried to exclude from your diet any foods that you like in order to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
4	Have you tried to follow definite rules regarding your eating (for example, a calorie limit) in order to influence your shape or weight (whether or not you have	0	1	2	3	4	5	6
5	Have you had a definite desire to have an empty stomach with the aim of influencing your shape or weight?	0	1	2	3	4	5	6
6	Have you had a definite desire to have a totally flat stomach?	0	1	2	3	4	5	6
7	Has thinking about food, eating or calories made it very difficult to concentrate on things you are interested in (for example, working, following a conversation or	0	1	2	3	4	5	6
8	Has thinking about shape or weight made it very difficult to concentrate on things you are interested in (for example, working, following a conversation or reading)?	0	1	2	3	4	5	6
9	Have you had a definite fear of losing control over eating?	0	1	2	3	4	5	6
10	Have you had a definite fear that you might gain weight?	0	1	2	3	4	5	6
11	Have you felt fat?	0	1	2	3	4	5	6

12	Have you had a strong desire to lose weight?	0	1	2	3	4	5	6

Questions 13-18: Please fill in the appropriate number in the boxes on the right. Remember that the questions only refer to the past four weeks (28 days).

Over the past four weeks (28 days)…….

13	Over the past 28 days, how many <u>times</u> have you eaten what other people would regard as an <u>unusually large amount of food</u> (given the circumstances)?	……
14	On how many of these times did you have a sense of having lost control over your eating (at the time that you were eating)?	……
15	Over the past 28 days, on how many <u>DAYS</u> have such episodes of overeating occurred (i.e. you have eaten an unusually large amount of food and have had a sense of loss of control at the time)?	……
16	Over the past 28 days, how many <u>times</u> have you made yourself sick (vomit) as a means of controlling your shape or weight?	……
17	Over the past 28 days, how many <u>times</u> have you taken laxatives as a means of controlling your shape or weight?	……
18	Over the past 28 days, how many times have you execised in a “driven” or “compulsive” way as a means of controlling your weight, shape or amount of fat or to burn off calories?	……

Questions 19-21: Please circle the appropriate number. Please note that for these questions the term “ binge eating” means eating what others would regard as an unusually large amount of food for the circumstances, accompanied by a sense of having lost control over eating.

	No days	1-5 days	6-12 days	13-15 days	16-22 days	23-27 days	Ever day
Over the past 28 days, on how many days have you eaten in secret (ie, furtively)?.....Do not count episodes of binge eating	0	1	2	3	4	5	6

		None of the times	A few of the times	Less than half	Half of the times	More than half	Most of the time	Every time
20	On what proportion of the times that you have eaten have you felt guilty (felt that you've done wrong) because of its effect on your shape or weight? .....Do not count episodes of binge eating	0	1	2	3	4	5	6

		Not at all		Slightly		Moderately		Markedly
21	Over the past 28 days, how concerned have you been about other people seeing you eat?							

Questions 22-28: Please circle the appropriate number on the right. Remember that the questions only refer to the past four weeks (28 days)

	On how many of the past 28 days ……	Not at all		Slightly		Moderately		Markedly
22	Has your <u>weight</u> influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
23	Has your <u>shape</u> influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
24	How much would it have upset you if you had been							
25	How dissatisfied have you been with your <u>weight</u>?	0	1	2	3	4	5	6
26	How dissatisfied have you been with your <u>shape</u>?	0	1	2	3	4	5	6
27	How uncomfortable have you felt seeing your body (for example, seeing your shape in the mirror, in a shop window reflection, while undressing or taking a bath or shower)?	0	1	2	3	4	5	6

28 How uncomfortable have you felt about others seeing your shape or figure (for example, in communal changing rooms, when swimming, or wearing tight clothes)?

What is your weight at present? (please give your best estimate)

What is your height? (please give your best estimate)

If female, over the past three-to-four months, have you missed any menstrual periods?

If so, how many?

Have you been taking the pill?

Appendix B. State-Trait Anxiety Inventory

State-Trait Anxiety Inventory for Adults

Self-Evaluation Questionnaire

STAI Form Y-1 and Form Y-2

Developed by Charles D. Spielberger
in collaboration with R.L. Gorsuch, R. Lushene, P.R. Vagg, and G.A. Jacobs

Published by Mind Garden
1690 Woodside Road Suite 202, Redwood City, CA 94061 USA 650-261-3500
www.mindgarden.com

SELF-EVALUATION QUESTIONNAIRE STAI Form Y-1

Please provide the following information:

Name_____Date_____S_____

Age_____ Gender (*Circle*) **M** **F** T_____

DIRECTIONS:

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you feel *right* now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

	NOT AT ALL	SOMEWHAT	MODERATELY SO	VERY MUCH SO
1. I feel calm	1	2	3	4
2. I feel secure	1	2	3	4
3. I am tense	1	2	3	4
4. I feel strained	1	2	3	4
5. I feel at ease	1	2	3	4
6. I feel upset	1	2	3	4
7. I am presently worrying over possible misfortunes	1	2	3	4
8. I feel satisfied	1	2	3	4
9. I feel frightened	1	2	3	4
10. I feel comfortable	1	2	3	4
11. I feel self-confident	1	2	3	4
12. I feel nervous	1	2	3	4
13. I am jittery	1	2	3	4
14. I feel indecisive	1	2	3	4
15. I am relaxed	1	2	3	4
16. I feel content	1	2	3	4
17. I am worried	1	2	3	4
18. I feel confused	1	2	3	4
19. I feel steady	1	2	3	4
20. I feel pleasant	1	2	3	4

STAIP-AD Test Form Y
www.mindgarden.com

SELF-EVALUATION QUESTIONNAIRE

STAI Form Y-2

Name_____ Date_____

DIRECTIONS

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you *generally* feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe how you generally feel.

	ALMOST NEVER	SOMETIMES	OFTEN	ALMOST ALWAYS
21. I feel pleasant	1	2	3	4
22. I feel nervous and restless	1	2	3	4
23. I feel satisfied with myself	1	2	3	4
24. I wish I could be as happy as others seem to be	1	2	3	4
25. I feel like a failure	1	2	3	4
26. I feel rested	1	2	3	4
27. I am "calm, cool, and collected"	1	2	3	4
28. I feel that difficulties are piling up so that I cannot overcome them	1	2	3	4
29. I worry too much over something that really doesn't matter	1	2	3	4
30. I am happy	1	2	3	4
31. I have disturbing thoughts	1	2	3	4
32. I lack self-confidence	1	2	3	4
33. I feel secure	1	2	3	4
34. I make decisions easily	1	2	3	4
35. I feel inadequate	1	2	3	4
36. I am content	1	2	3	4
37. Some unimportant thought runs through my mind and bothers me	1	2	3	4
38. I take disappointments so keenly that I can't put them out of my mind	1	2	3	4
39. I am a steady person	1	2	3	4
40. I get in a state of tension or turmoil as I think over my recent concerns and interests	1	2	3	4

STAIP-AD Test Form Y
www.mindgarden.com

Appendix C. Yale Brown Obsessive Compulsive Scale

YALE-BROWN OBSESSIVE COMPULSIVE SCALE (Y-BOCS)

General Instructions

This rating scale is designed to rate the severity and type of symptoms in patients with obsessive compulsive disorder (OCD). In general, the items depend on the patient's report; however, the final rating is based on the clinical judgment of the interviewer. Rate the characteristics of each item during the prior week up until and including the time of the interview. Scores should reflect the average (mean) occurrence of each item for the entire week.

This rating scale is intended for use as a semi-structured interview. The interviewer should assess the items in the listed order and use the questions provided. However, the interviewer is free to ask additional questions for purposes of clarification. If the patient volunteers information at any time during the interview, that information will be considered. Ratings should be based primarily on reports and observations gained during the interview. If you judge that the information being provided is grossly inaccurate, then the reliability of the patient is in doubt and should be noted accordingly at the cad of the interview (item 19).

Additional information supplied by others (e.g., spouse or parent) may be included in a determination of the ratings only if it is judged that (1) such information is essential to adequately assessing symptom severity and

(2) consistent week-to-week reporting can be ensured by having the same informant(s) present for each rating session.

Before proceeding with the questions, define "obsessions" and "compulsions" for the patient as follows:

"OBSESSIONS are unwelcome and distressing ideas, thoughts, images or impulses that repeatedly enter your mind. They may seem to occur against your will. They may be repugnant to you, you may recognize them as senseless, and they may not fit your personality."

"COMPULSIONS, on the other hand, are behaviors or acts that you feel driven to perform although you may recognize them as senseless or excessive. At times, you may try to resist doing them but this may prove difficult. You may experience anxiety that does not diminish until the behavior is completed."

Let me give you some examples of obsessions and compulsions."

"An example of an obsession is: the recurrent thought or impulse to do serious physical harm to your children even though you never would."

"An example of a compulsion is: the need to repeatedly check appliances, water faucets, and the lock on the front door before you can leave the house. While most compulsions are observable behaviors, some are unobservable mental acts, such as silent checking or having to recite nonsense phrases to yourself each time you have a bad thought."

"Do you have any questions about what these words mean?" [If not, proceed.]

On repeated testing it is not always necessary to re-read these definitions and examples as long as it can be established that the patient understands them. It may be sufficient to

remind the patient that obsessions are the thoughts or concerns and compulsions are the things you feel driven to do, including covert mental acts.

Have the patient enumerate current obsessions and compulsions in order to generate a list of target symptoms. Use the Y-BOCS Symptom Checklist as an aid for identifying current symptoms. It is also useful to identify and be aware of past symptoms since they may re-appear during subsequent ratings. Once the current types of obsessions and compulsions are identified, organize and list them on the Target Symptoms form according to clinically convenient distinctions (e.g., divide target compulsions into checking and washing). Describe salient features of the symptoms so that they can be more easily tracked (e.g., in addition to listing checking, specify what the patient checks for). Be sure to indicate which symptoms are the most prominent i.e., those that will be the major focus of assessment. Note, however, that the final score for each item should reflect a composite rating of all of the patient's obsessions or compulsions.

The rater must ascertain whether reported behaviors are bona fide symptoms of OCD and not symptoms of another disorder, such as Simple Phobia or a Paraphilia. The differential diagnosis between certain complex motor tics and certain compulsions (e.g., involving touching) may be difficult or impossible. In such cases, it is particularly important to provide explicit descriptions of the target symptoms and to be consistent in subsequent ratings. Separate assessment of tic severity with a tic rating instrument may be necessary in such cases. Some of the items listed on the Y-BOCS Symptom Checklist, such as trichotillomania, are currently classified in DSM-m-R as symptoms of an Impulse Control Disorder. It should be noted that the suitability of the Y-BOCS for use in disorders other than DSM-m-R-defined OCD has yet to be established. However, when using the Y-BOCS to rate severity of symptoms not strictly classified under OCD (e.g., trichotillomania) in a

patient who otherwise meets criteria for OCD, it has been our practice to administer the Y-BOCS twice: once for conventional obsessive-compulsive symptoms, and a second time for putative OCD-related phenomena. In this fashion separate Y-BOCS scores are generated for severity of OCD and severity of other symptoms in which the relationship to OCD is still unsettled.

On repeated testing, review and, if necessary, revise target obsessions prior to rating item 1. Do likewise for compulsions prior to rating item 6. All 19 items are rated, but only items 1-10 (excluding items 1b and 6b) are used to determine the total score. The total Y-BOCS score is the sum of items 1-10 (excluding 1b and 6b), whereas the obsession and compulsion subtotals are the sums of items 1-5 (excluding lb) and 10 (excluding 6b3; respectively. Because at the time of this writing (9/89) there are limited data regarding the psychometric properties of items 1b, 6b, and 11-16, these items should be considered investigational. Until adequate studies of reliability, validity, and sensitivity to change of those items are conducted, we must caution against placing much weight on results derived from these item scores. These important caveats aside, we believe that items lb (obsession-free interval), 6b (compulsion-free interval), and 12 (avoidance) may provide information that has bearing on the severity of obsessive-compulsive symptoms. Item 11 (insight) may also furnish useful clinical information. We are least secure about the usefulness of items 13-16. Items 17 (global severity) and 18 (global improvement) have been adapted from the Clinical Global Impression Seale (Guy W, 1976) to provide measures of overall functional impairment associated with, but not restricted to, the presence of obsessive-compulsive symptoms. Disability produced by secondary depressive symptoms would also be considered when rating these items. Item 19, which estimates the reliability of the information reported by the patient, may assist in the interpretation of scores on other Y-

BOCS items in some cases of OCD. YALE-BROWN OBSESSIVE COMPULSIVE SCALE (Y-BOCS) www.cnsforum.com 4

Y-BOCS SYMPTOM CHECKLIST (9/89)

Check all that apply, but clearly mark the principal symptoms with a "P", (Rater must ascertain whether reported behaviors are bona fide symptoms of OCD, and not symptoms of another disorder such as Simple Phobia or Hypochondriasis. Items marked "*" may or may not be OCD phenomena.)

AGGRESSIVE OBSESSIONS

	Current	Past	Examples
1. Fear might harm self			Fear of eating with a knife or fork, fear of handling sharp objects, fear of walking near glass windows.
2. Fear might harm others			Fear of poisoning other people's food, fear of harming babies, fear of pushing someone in front of a train, fear of hurting someone's feelings, fear of being responsible by not providing assistance for some imagined catastrophe, fear of causing harm by bad advice.
3. Violent or horrific images			Images of murders, dismembered bodies, or other disgusting scenes.
4. Fear of blurting out obscenities or insults			Fear of shouting obscenities in public situations like church, fear of writing obscenities.
5. Fear of doing something else embarrassing *			Fear of appearing foolish in social situations
6. Fear will act on unwanted impulses			Fear of driving a car into a tree, fear of running someone over, fear of stabbing a friend.
7. Fear will steal things			Fear of "cheating" a cashier, fear of shoplifting inexpensive items.
8. Fear will harm others because not careful enough			Fear of causing an accident without being aware of it (such as a hit-and-run automobile accident).
9. Fear will be responsible for something else terrible			Fear of causing a fire or burglary because of not being careful enough in checking the house before leaving.

happening			
10. Other:_____			

CONTAMINATION OBSESSIONS

	Current	Past	Examples
11. Concerns or disgust with bodily waste or secretions.			Fear of contracting AIDS, cancer, or other diseases from public rest rooms; fears of your own saliva, urine, feces, semen, or vaginal secretions.
12. Concern with dirt or germs.			Fear of picking up germs from sitting in certain chairs, shaking hands, or toughing door handles.
13. Excessive concern with environmental contaminants.			Fear of being contaminated by asbestos or radon, fear of radioactive substances, fear of things associated with towns containing toxic waste sights.
14. Excessive concern with household items			Fear of poisonous kitchen or bathroom cleansers, solvents, Insect spray or turpentine.
15. Excessive concern with animals.			Fear of being contaminated by touching an insect, dog, cat, or other animal.
16. Bothered by sticky substances or residues			Fear of adhesive tape or other sticky substances that may trap contaminants.
17. Concerned I will get ill because of contaminant			Fear of getting ill as a direct result of being contaminated (beliefs vary about how long the disease will take to appear.
18. Concerned I will get others ill by spreading contaminant (Aggressive)			Fear of touching other people or preparing their food after you touch poisonous substances (like gasoline) or after you touch your own body.
19. Other:_____			

SEXUAL OBSESSIONS

	Current	Past	Examples
20. I have forbidden or perverse sexual thoughts, images, or impulses			Unwanted sexual thoughts about strangers, family, or friends.
21. Content involves children or incest			Unwanted thoughts about sexually molesting either your own children or other children.
22. Content involves homosexuality			Worries like “Am I a homosexual?” or “What if I

*			suddenly become gay?" when there is no basis for these thoughts.
23. Aggressive sexual behavior toward others. *			Unwanted images of violent sexual behavior toward adult strangers, friends, or family members.
24. Other:_____			

Hoarding / Saving Obsessions

Distinguish from hobbies and concern with objects of monetary or sentimental value.	**Current**	**Past**	**Examples**
25. I have obsessions about hoarding or saving things.			Worries about throwing away seemingly unimportant things that you might need in the future, urges to pick up and collect useless things.

RELIGIOUS OBSESSIONS

	Current	**Past**	**Examples**
26. (Scrupulosity) Concerned with sacrilege and blasphemy			Worries about having blasphemous thoughts, saying blasphemous things, or being punished for such things.
27. Excess concern with right/wrong, morality.			Worries abut always doing "the right thing", having told a lie, or having cheated someone.
28. Other:_____			

OBSESSION WITH NEED FOR SYMMETRY OR EXACTNESS

	Current	**Past**	**Examples**
29. (Accompanied by magical thinking (c.x., concerned the mother will have an accident unless things are in the right place) Obsessions about symmetry or exactness			Worries about papers and books being properly aligned, worries about calculations or handwriting being perfect.
30. Not accompanied by magical thinking			

MISCELLANEOUS OBSESSIONS

	Current	**Past**	**Examples**
31. Need to know or remember certain things			Belief that you need to remember insignificant things like license plate numbers, the names of actors on television shows, old telephone numbers, bumper

			sticker or t-shirt slogans.
32. Fear of saying certain things			Fear of saying certain words (such as “thirteen”) because of superstitions, fear of saying something tat might be disrespectful to a dead person, fear of using words with an apostrophe (because this denotes possession).
33. Fear of not saying just the right thing			Fear of having said the wrong thing, fear of not using the “perfect” word.
34. Fear of losing things			Worries about losing a wallet or other unimportant objects, like a scrap of note paper.
35. Intrusive (non-violent) images			Random unwanted images in your mind.
36. Intrusive nonsense sounds, words, or music.			Words, songs, or music in your mind that you can’t stop.
37. Bothered by certain sounds/noises *			Worries about the sounds of clocks ticking loudly or voices in another room that may interfere with sleeping.
38. Lucky/unlucky numbers			Worries about common numbers (like thirteen) that may cause you to perform activities a certain number of times or to postpone an action until a certain lucky hour of the day.
39. Colors with special significance			Fear of using objects of certain colors (e.g. black may be associated with death, red with blood and injury).
40. Superstitious fears			Fear of passing a cemetery, hearse, or black cat; fear of omens associated with death.

SOMATIC OBSESSIONS

	Current	**Past**	**Examples**
55. Concern with illness or disease *			Worries that you have an illness like cancer, heart disease or AIDS, despite reassurance from doctors that you do not.
41. Excessive concern with body part or aspect of appearance (e.g. dysmorphophobia) *			Worries that your face, ears, nose, eyes, or another part of your body is hideous, ugly, despite reassurances to the contrary.
42. Other _____			

CLEANING/WASHING COMPULSIONS

	Current	Past	Examples
43. Excessive or ritualized hand washing			Washing your hands many times a day or for long periods of time after touching, or thinking that you have touched a contaminated object. This may include washing the entire length of your arms.
44. Excessive or ritualized showering, bathing, tooth brushing, grooming, or toilet routine.			Taking showers or baths or performing other bathroom routines that may last for several hours. If the sequence is interrupted the entire process may have to be restarted.
45. Excessive or ritualized cleaning of household items or other inanimate objects			Excessive cleaning of faucets, toilets, floors, kitchen counters, or kitchen utensils.
46. Other measures to prevent or remove contact with contaminants			Asking family members to handle or remove insecticides, garbage, gasoline cans, raw meat, paints, varnish, drugs in the medicine cabinet, or kitty litter. If you can't avoid these things, you may wear gloves to handle them, such as when using a self-service gasoline pump.
47. Other _____			

CHECKING COMPULSIONS

	Current	Past	Examples
48. Checking locks, stove, appliances, etc.			Washing your hands many times a day or for long periods of time after touching, or thinking that you have touched a contaminated object. This may include washing the entire length of your arms.
49. Checking that did not/will not harm others.			Checking that you haven't hurt someone without knowing it. You may ask other for reassurance or telephone to make sure that everything is all right
50. Checking that did not/will not harm self			Looking for injuries or bleeding after handling sharp or breakable objects. You may frequently go to doctors to ask for reassurance that you haven't hurt yourself.
51. Checking that nothing terrible did/will happen			Searching the newspaper or listening to the radio or television for news about some catastrophe that you believe you caused. You may also ask people for

			reassurance that you didn't cause an accident.
52. Checking that did not make mistake			Repeated checking of door locks, stoves, electrical outlets, before leaving home; repeated checking while reading, writing, or doing simple calculations to make sure that you didn't make a mistake (you can't be certain that you didn't).
53. Checking tied to somatic obsessions			Seeking reassurance from friends or doctors that you aren't having a heart attack or getting cancer; repeatedly taking your pulse, blood pressure, or temperature; checking yourself for body odors; checking your appearance in a mirror, looking for ugly features.
54. Other _____			

REPEATING COMPULSIONS

	Current	**Past**	**Examples**
55. Re-reading or re-writing			Taking hours to read a few pages in a book or to write a short letter because you get caught in a cycle of reading and rereading; worrying that you didn't understand something you just read; searching for a "perfect" word or phrase; having obsessive thoughts about the shape of certain printed letters in a book.
56. Need to repeat routine activities			Repeating activities like turning appliances on and off, combing your hair, going in and out of a doorway, or looking in a particular direction; not feeling comfortable unless you do these things the "right" number of times.
57. Other			

COUNTING COMPULSIONS

	Current	**Past**	**Examples**
58. Need to count and recount			Counting objects like ceiling or floor tiles, books in a bookcase, nails in a wall, or even grains of sand on a beach; counting when you repeat certain activities, like washing.

ORDERING / ARRANGING COMPULSIONS

	Current	Past	Examples
59. Need to order and reorder, arrange and rearrange items.			Straightening paper and pens on a desktop or books in a bookcase, sating hours arranging things in your house in “order” and then becoming very upset if this order is disturbed.

HOARDING / COLLECTING COMPULSIONS

Distinguish from hobbies and concern with objects of monetary or sentimental value.	Current	Past	Examples
60. Compulsions to hoard or collect things.			Saving old newspapers, notes, cans, paper towels, wrappers, and empty bottles for fear that if you throw them away you may one day need them; picking up useless objects from the street or from the garbage can.

MISCELLANEOUS COMPULSIONS

Mental rituals other than checking / counting.	Current	Past	Examples
61. Mental rituals (other than checking / counting).			Performing rituals in your head, like saying prayers or thinking a “good” thought to undo a “bad” thought. These are different from obsessions because you perform them intentionally to reduce anxiety or to feel better.
62. Need to tell, ask, or confess			Asking other people to reassure you, confessing to wrong behaviors you never even did, believing that you have to tell other people certain words to feel better.
63. Need to touch, tap, or rub *			Giving in to the urge to touch rough surfaces, like wood, or hot surfaces, like a stove top; giving in to the urge to lightly touch other people; believing you need to touch an object like a telephone to prevent an illness in your family.
64. Measures (not checking) to prevent harm or terrible consequences to myself or others.			Staying away from sharp or breakable objects, such as knives, scissors, and fragile glass.

65. Ritualized eating behaviors *			Arranging your food, knife, and fork in a particular order before being able to eat, eating according to a strict ritual, not being able to eat until the hands of a clock point exactly to a certain time.
66. Superstitious behaviors			Not taking a bus or train if its number contains an “unlucky” number (like thirteen), staying in your house on the thirteenth of the month, throwing away clothes you wore while passing a funeral home or cemetery.
67. Hair pulling, Trichotillomania *			Pulling hair from your scalp, eyelids, eyelashes, or pubic areas, using your fingers or tweezers. You may produce bald spot that require you to wear a wig, or you may pluck your eyebrows or eyelids smooth.

YALE-BROWN OBSESSIVE COMPULSIVE SCALE (Y-BOCS)

"I am now going to ask several questions about your obsessive thoughts." [Make specific reference to the patient's target obsessions.]

1. TIME OCCUPED BY OBSESSIVE THOUGHTS

0 = None.
1 = Mild, less than 1 hr/day or occasional intrusion.
2 = Moderate, 1 to 3 hrs/day or frequent intrusion.
3 = Severe, greater than 3 and up to 8 hrs/day or very frequent intrusion.
4 = Extreme, greater than 8 hrs/day or near constant intrusion.

Q: How much of your time is occupied by obsessive thoughts? [When obsessions occur as brief, intermittent intrusions, it may be difficult to assess time occupied by them in temns of total hours. In such cases, estirnate time by detesmining how frequently they occur. Consider both the number of times the intrusions occur and how many hours of the day are affected. Ask:1 How frequently do the obsessive thoughts occur? [Be sure to exclude ruminations and preoccupations which, unlike obsessions, are ego-syntonic and rational (but exaggerated).]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

I b. OBSESSION-FREE INTERVAL (not included in total score)

0 = No symptoms.

1 = Long symptom-free interval, more than 8 consecutive hours/day symptom-free.

2 = Moderately long symptom-free interval, more than 3 and up to 8 consecutive hours/day symptom-free.

3 = Short symptom-free interval, from I to 3 consecutive hours/day symptom-free.

4 = Extremely short symptom-free interval, less than I consecutive hour/day symptom-free.

Q: On the average, what is the longest number of consecutive waking hours per day that you are completely free of obsessive thoughts? [If necessary, ask:1 What is the longest block of time in which obsessive thoughts are absent?	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

2. INTERFERENCE DUE TO OBSESSIVE THOUGHTS

0 = None.

1 = Mild, slight interference with social or occupational activities, but overall performance not impaired.

2 = Moderate, definite interference with social or occupational performance, but still manageable.

3 = Severe, causes substantial impairment in social or occupational performance.

4 - Extreme, incapacitating.

Q: How much do your obsessive thoughts interfere with your social or work (or role) functioning? Is there anything that you don't do because of them? [If currently not working determine how much performance would be affected if patient were employed.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

3. DISTRESS ASSOCIATED W1TH OBSESSIVE THOUGHTS

0 = None

I = Mild, not too disturbing

2 = 1doderate, disturbing, but still manageable

3 = Severe, very disturbing

4 = Extreme, near constant and disabling distress

Q: How much distress do your obsessive thoughts cause you? [In most eases, distress is equated with anxiety; however, patients may report that their obsessions are "disturbing" but deny "anxiety." Only rate anxiety that seems triggered by obsessions, not generalized anxiety or associated with other conditions.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

4. RESISTANCE AGAINST OBSESSIONS

0 = Makes an effort to always resist, or symptoms so minimal doesn't need to actively resist

1 = Tries to resist most of the time

2 = Makes some effort to resist

3 = Yields to all obsessions without attempting to control them, but does so with some reluctance

4 = Completely and willingly yields to all obsessions

Q: How much of an effort do you make to resist the obsessive thoughts? How often do you try to disregard or turn your anention away from these thoughts as they eater your mind? [Only rate effort made to resist, not success or failure in actually controlling the obsessions. How much the patient resists the obsessions may or may not correlate with his/her abilig to control them. Note that this item does not direetly measure the severig of the intrusive thoughts; rather it rates a manifestation of health, i.e., the effort the patient makes to counteract the obsessions by means other than avoidance or the performance of compulsions. Thus, the more the patient tries to resist, the less impaired is this aspect of his/her functioning. There are "active" and "passive" forms of resistance. Patients in behavioral therapy may be encouraged to counteract their obsessive symptoms by not struggling against them (e.g., "just let the thoughts come; passive opposition) or by intentionally bringing on the disturbing thoughts. For the purposes of this item, consider use of these behavioral techniques as forms of resistance. If the obsessions are minimal, the patieut may not feel the need to resist them. In such cases, a rating of "0" should be given.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

5. DEGREE OF CONTROL OVER OBSESSIVE THOUGHTS

0 = Complete control.

1 = Much control, usually able to stop or divert obsessions with some effort and concentration.

2 = Moderate control, sometimes able to stop or divert obsessions.

3 = Little control, rarely successful in stopping or dismissing obsessions, can only divert attention with difficulty.

4 = No control, experienced as completely involuntary, rarely able to even momentarily alter obsessive thinking.

Q: How much control do you have over your obsessive thoughts? How successful are you in stopping or diverting your obsessive thinking? Can you dismiss them? [In contrast to the preceding item on resistance, the ability of the patient to control his obsessions is more closely related to the severity of the intrusive thoughts.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

"The next several questions are about your compulsive behaviors."
[Make specific reference to the patient's target compulsions.]

6. TIME SPENT PERFORM~G COMPULSIVE BEHAVIORS

0 = None

1 = Mild (spends less than I hr/day performing compulsions), or occasional performance of compulsive behaviors.

2 = Moderate (speeds from I to 3 hrs/day performing compulsions), or frequent performance of compulsive behaviors.

3 = Severe (spends more than 3 and up to 8 hrs/day performing compulsions), or very frequent performance of compulsive behaviors.

4 = Extreme (spends more than 8 hrs/day performing compulsions), or near constant performance of compulsive behaviors (too numerous to count).

Q: How much time do you spend performing compulsive behaviors? [When rituals involving activities of daily living are chiefly present, ask:] How much longer than most people does it take to complete routine activities because of your rituals? [When compulsions occur as brief, intermittent behaviors, it may difficult to assess time spent performing them in terms of total hours. In such cases, estimate time by determining how frequently they are performed. Consider both the number of times compulsions are performed and how many hours of the day are affected. Count separate occurrences of compulsive behaviors, not number of repetitions; e.g., a patient who goes into the bathroom 20 different times a day to wash his hands 5 times very quickly, performs compulsions 20 times a day, not 5 or 5 x 20 = 100. Ask:] How frequently do you perform compulsions? 1In most cases compulsions are observable behaviors(e.g., land washing), but some compulsions are covert (e.g., silent checking).]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

6b. COMPULSION-FREE INTERVAL(not included in total score)

0 = No symptoms.

1 = Long symptom-free interval, more than 8 consecutive hours/day symptom-free.

2 = Moderately long symptom-free interval, more than 3 and up to 8 consecutive hours/day symptom-free.

3 = Short symptom-free interval, from I to 3 consecutive hours/day symptom-free.

4 = Extremely short symptom-free interval, less than I consecutive hour/day symptom-free.

Q: On the average, what is the longest number of consecutive waking hours per day that you are completely free of compulsive behavior? [If necessary, ask:] What is the longest block of time in which compulsions are absent?different times a day to wash his hands 5 times very quickly, performs compulsions 20 times a day, not 5 or 5 x 20 = 100. Ask:] How frequently do you perform compulsions? 1In most cases compulsions are observable behaviors(e.g., land washing), but some compulsions are covert (e.g., silent checking).]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

7 INTERFERIINCE DUE TO COMPULSIVE BEHAVIQRS

0 = None

1 = Mild, slight interference with social or occupational activities, but overall performance not impaired

2 = Moderate, definite interference with social or occupational performance, but still manageable

3 = Severe, causes substantial impaiment in social or occupational performance

4 = Extreme, incapacitating

Q: How much do your compulsive behaviors interfere with your social or work (or role) functioning? Is there anything that you don't do because of the compulsions? [If currently not working determine how much performance would be affected if patient were employed.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

8. DISTRESS ASSOCIATED WITH COMPULSIVE BEHAVIOR

0 = None

I = Mild only slightly anxious if compulsions prevented, or only slight anxiety during performance of compulsions

2 = Moderate, reports that anxioty would mount but remain manageable if compulsions prevented, or that anxiety increases but remains manageable during performance of compulsions

3 = Severe, prominent and very disturbing increase in anxiety if compulsions interrupted, or prorninent and very disturbing increase in anxiety during performance of compulsions

4 = Extreme, incapacitating anxiety from any intervention aimed at modifying activity, or incapacitating anxiety develops during performance of compulsions

Q: How would you feel if prevented from performing your compulsion(s)? [Pause] How anxious would you become? [Rate degree of distress patient would experience if performance of the compulsion were suddenly interrupted without reassurance offered. In most, but not all cases, performing compulsions reduces anxiety. If, in the judgement of the interviewer, anxiety is actually reduced by preventing compulsions in the manner described above, then asked: How anxious do you get while performing compulsions until you are satisfied they are completed?	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

9. RESISTANCE AGAINST COMPULSIONS

0 = Malces an effort to always resist, or symptoms so minimal doesn't need to actively resist

I = Tries to resist most of the time

2 = Makes some effort to resist

3 = Yields to almost all compulsions without attempting to control them, but does so with some reluetance

4 = Completely and willingly yields to all compulsions

Q: How much of an effort do you make to resist the compulsions? I Only rate effort made to resist, not success or failure in actually controlling the compulsions. How much the patient resists the compulsions may or may not correlate with his ability to control them. Note that this item does not directly measure the severity of the compulsions; rather it rates a manifestation of health, i.e., the effort the patient makes to counteract the compulsions. Thus, the more the patient tries to resist, the less impaired is this aspect of his functioning. If the compulsions are minimal, the patient may not feel the need to resist them. In such cases, a rating of "0" should be given.]	0 1 2 3 4

10. DEGREE OF CONTROL OVER COMULSIVE BEHAVIOR

I = Much control, experiences pressure to perform the behavior but usually able to exercise voluntary control over it.

2 = Moderate control, strong pressure to perform behavior, can control it only with difficulty

3 = Little control, very strong drive to perform behavior, must be carried to completion, can only delay with difficulty

4 = No control. drive to perform behavior expericoced as completely involuntary and overpowering, rarely able to even momentarily delay activity

Q: How strong is the drive to perform the compulsive behavior? [Pause] How much control do you have over the compulsions? [In contrast to the preceding item on resistance, the ability of the patient to control his compulsions is more closely related to the severity of the compulsions.]	0 1 2 3 4

"The remaining questions are about both obsessions and compulsions. Some ask about related problems." These are investigational items not included in total Y-BOCS score but may be useful in assessing these symptoms.

11. INSIGHT INTO OBSESSIONS AND COMPULSIONS

0 = Excellent insight, fully rational

1 = Good insight. Readily acknowledges absurdity or excessiveness of thoughts or behaviors but does not seem completely convinced that there isn't something besides anxiety to be concerned about (i.e., has lingering doubts).

2 = Fair insight. Reluctantly admits thoughts or behavior seem unreasonable or excessive, but wavers. May have some unrealistic fears, but no fixed convictions.

3 = Poor insight. Maintains that thoughts or behaviors are not unreasonable or excessive, but acknowledges validity of contrary evidence (i.e., overvalued ideas present).

4 = Lacks insight, delusional. Definitely convinced that concerns and behavior are reasonable, unresponsive to contrary evidence.

Q: Do you think your concerns or behaviors are reasonable? [Pause] What do you think would happen if you did not perform the compulsion(s)? Are you convinced something would really happen? 1Ratc patient's insight into the senselessness or excessiveness of his obsession(s) based on beliefs expressed at the time of the interview.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

12. AVOIDANCE

0 = No deliberate avoidance

1 = Mild, minimal avoidance

2 = Moderate, some avoidance; clearly present

3 = Severe, much avoidance; avoidance prominent

4 = Extreme, very extensive avoidance; patient does almost everything he/she can to avoid triggering symptoms

Q: Have you been avoiding doing anything, going any place, or being with anyone because of your obsessional thoughts or out of concern you will perform compulsions? [If yes, then ask:] Elow much do you avoid? [Rate degree to which patient deliberately tries to avoid things. Sometimes compulsions are designed to "avoid" contact with something that the patient fears. For example, clothes washing rituals would be designated as compulsions, not as avoidant behavior. If the patient stopped doing the laundry then this would constitute avoidance.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

13. DEGEE OF INDECISIVENESS

0 = None

1 = Mild, some trouble making decisions about minor things

2 = Moderate, freely reports significant trouble making decisions that others would not think twice about

3 = Severe, continual weighing of pros and cons about nonessentials.

4 = Extreme, unable to make any decisions. Disabling.

Q: Do you have trouble making decisions about little things that other people might not think twice about (e.g., which clothes to put on in the morning; which brand of cereal to buy)? [Exclude difficulty making decisions which reflect ruminative thinking. Ambivalence concerning rationally-based difficult choices should also be excluded.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

14. OVERVALUED SENSE OF RESPONSIBILY

0 = None I = Mild, only mentioned on questioning, slight sense of over-responsibility

2 = Moderate, ideas stated spontaneously, clearly present; patient experiences significant sense of over-responsibility for events outside his/her reasonable control

3 = Severe, ideas prominent and pervasive; deeply concerned he/she is responsible for events clearly outside his control. Self-blaming farfetched and nearly irrational

4 = Extreme, delusional sense of responsibility (e.g., if an earthquake occurs 3,000 miles away patient blames herself because she didn't perform her compulsions)

Q: Do you feel very responsible for the consequences of your actions? Do you blame yourself for the outcome of events not completely in your control? [Distinguish from normal feelings of responsibility, feelings of worthlessness, and pathological guilt. A guilt-ridden person experiences himself or his actions as bad or evil.]	0 1 2 3 4

15. PERVASIVE SLOWNESS/ DISTURBANCE OF INERTIA

0 = None.

I = Mild, occasional delay in starting or finishing.

2 = Moderate, frequent prolongation of routine activities but tasks usually completed. Frequently late.

3 = Severe, pervasive and marked difficulty initiating and completing routine tasks. Usually late.

4 = Extreme, unable to start or complete routine tasks without full assistance.

Q: Do you have difficulty starting or finishing tasks? Do many routine activities take longer than they should? [Distinguish from psychomotor retardation secondary to depression. Rate increased time spent performing routine activities even when specific obsessions cannot be identified.]	0 1 2 3 4

16. PATHOLOGICAL DOUBTING

0 = None.

1 = Mild, only mentioned on questioning, slight pathological doubt. Examples given may be within normal range.

2 = Moderate, ideas stated spontaneously, clearly present and apparent in some of patient's behaviors, patient bothered by significant pathological doubt. Some effect on performance but still manageable.

3 = Severe, uncertainty about perceptions or ,memory prominent; pathological doubt frequently affects performance.

4 = Extreme uncertainty about perceptions constantly present; pathological doubt substantially affects almost all activities. Incapacitating (e.g., patient states "my mind doesn't trust what my eyes see").

Q: After you complete an activity do you doubt whether you performed it correctly? Do you doubt whether you did it at all? When carrying out routine activities do you find that you don't trust your senses (i.e., what you see, hear, or touch)?	0 1 2 3 4

[Items 17 and 18 refer to global illness severity. The rater is required to consider global function, not just the severity of obssive-compulsive symptoms.]

17. GLOBAL SEVERITY:

0 = No illness
1 = Illness slight, doubtful, transient; no functional impairment
2 = Mild symptoms, little functional impairment
3 = Moderate symptoms, functions with effort
4 = Moderate - Severe symptoms, limited functioning
5 = Severe symptoms, functions mainly with assistance
6 = Extremely Severe symptoms, completely nonfunctional

Interviewer's judgement of the overall severity of the patient's illness. Rated from O (no illness) to 6-(most severe patient seen). [Consider the degree of distress reported by the patient, the symptoms observed, and the functional impairment reported. Your judgement is required both in averaging this data as well as weighing the reliability or accuracy of the data obtained. This judgement is based on information obtained during the interview.]	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6

18. GLOBAL IMPROVEMENT:

0 = Very much worse
1 = Much worse
2 = Minimal worse
3 = No change
4 = Minimally improved
5 = Much improved
6 = Very much improved

Rate total overall improvement present SINCE THE INITIAL RATING whether or not, in your judgement, it is due to drug treatment.	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6

19. RELIABILITY:

0 = Excellent, no reason to suspect data unreliable . .
1 = Good, factor(s) present that may adversely affect reliability
2 = Fair, factorts) present that definitely reduce reliability
3 = Poor, very low reliability

Rate the overall reliability of the rating scores obtained. Factors that may affect reliability include the patient's cooperativenes and his/her natural ability to communicate. The type and severity of obsessive-compulsive symptoms present may interfere with the patient's concentration, attention, or freedom to speak spontaneously (e.g., the content of some obsessions may cause the patient to choose his words very carefully).	☐ 0 ☐ 1 ☐ 2 ☐ 3

Items 17 and 18 arc adapted from the Clinical Global Impression Scale (Guy W: ECDEU Assessment Manual for Psychopharrnacology: Publication 76-338. Washington, D.C., U.S. Department of Health, Education, and Welfare (1976)).

Additional infomnation regarding the development, use, and psychometric properties of the Y-BOCS can be found in Goodman WK, Price LH, Rasmussen SA, et al.: The Yale-Brown Obsessive Compulsive Scaie (YBOCS): Part I. Development, use, and reliability. Arch Gen Psvchiaty (46:1006~1011, 1989). and Goodman WK, Price LH, Rasmussen SA, ct al.: The Yale-Brown Obsessive Compulsive Scale (Y-BOCS): Part II. Validity. Arch Gen Psvchiatry (46:1012-1016, 1989).

Copies of a version of the Y-BOCS modified for usc in children, the Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS) (Goodman WK, Rasmussen SA, Price LH, Mazure C, Rapoport JL, Heninger GR, Charney DS), is available from Dr. Goodman on request.

Appendix D. Children's Obsessional Compulsive Inventory

CHOCI: Part 1

To be completed by the child/adolescent

Each of the following questions asks you about things or "habits" you feel you have to do although you may know that they do not make sense. Sometimes, you may try to stop from doing them but this might not be possible. You might feel worried or angry or frustrated until you have finished what you have to do. An example of a habit like this may be the need to wash your hands over and over again even though they are not really dirty, or the need to count up to a special number (e.g., 6 or 10) while you do certain things.

Please answer each question by putting a circle around the number that best describes how much you agree with the statement, or how much you think it is true of you. Please answer each item, without spending too much time on any one item. There are no right or wrong answers.

Example:	**Not at all**	**Somewhat**	**A lot**
I feel that I must check and check again that the stove is turned off, even if I don't want to do so.	**1**	**2**	**3**

How much do you agree with each of the following statements?	**Not at all**	**Somewhat**	**A lot**
1. I spend far too much time washing my hands over and over again.	1	2	3
2. I feel I must check my homework over and over and over again.	1	2	3
3. I feel I must do ordinary/everyday things exactly the same way, every time I do them.	1	2	3
4. I find it very difficult to throw things away so my room is much too crowded/cluttered.	1	2	3
5. I am too scared to use public toilets because of dirt/germs.	1	2	3
6. I check and check things like taps and switches over and over again.	1	2	3
7. I get really upset if my things are not always in exactly the same place.	1	2	3

How much do you agree with each of the following statements?	Not at all	Somewhat	A lot
8. I often get behind in my school work because I write the same words over and over and over again.	1	2	3
9. I am much, much too concerned about being clean.	1	2	3
10. I spend a lot of time every day checking things over and over and over again.	1	2	3
11. I often have trouble finishing things because I need to make absolutely sure that everything is exactly right.	1	2	3
12. I spend far too much time arranging my things in order.	1	2	3
13. I need someone to tell me things are alright over and over again.	1	2	3
14. I find it very, very upsetting to touch garbage or garbage bins.	1	2	3
15. I check and check over and over again that my doors or windows are locked, even though I try not to do so.	1	2	3
16. I always feel I must get dressed in exactly the same order every single day.	1	2	3
17. I always count, even when doing ordinary things.	1	2	3
18. I often feel I have to ask someone the same question over and over and over again.	1	2	3
19. I am often very late because I can't finish things on time.	1	2	3

Please try to think about the three most upsetting **habits** that you feel you **have** to do and **can't stop**. For example, feeling that you have to wash your hands far too often, or check that your school work is just right.

1)_____

2)_____

3)_____

How much time do you spend doing these habits? Please circle the answer that best describes you.

0	1	2	3	4
None	Less than 1 hr. a day (occassionally)	1-3 hrs. a day (part of a morning or afternoon)	3-8 hrs. a day (about half the time you're awake)	More than 8 hrs. a day (almost all the time you're awake)

How much do these habits get in the way of school or doing things with friends? Please circle the answer that best describes you.

0	1	2	3	4
Not at all	A little	Somewhat	A lot	Almost always

How would you feel if prevented from carrying out your habits? How upset would you become? Please circle the answer that best describes you.

0	1	2	3	4
Not at all	A little	Somewhat	A lot	Totally

How much do you try to fight the upsetting habits? Please circle the answer that best describes you.

0	1	2	3	4
I always try to resist	I try to resist most of the time	I make some effort to resist	Even though I want to, I don't try to resist	I don't resist at all

How strong is the feeling that you have to carry out the habits? Please circle the answer that best describes you.

0	1	2	3	4
Not strong	Mild pressure to carry out habits	Strong pressure to carry out habits; hard to control	Very strong pressure to carry out habits; very hard to control	Extreme pressure to carry out habits; impossible to control

How much have you been avoiding doing anything, going any place, or being with anyone because of your upsetting habits? Please circle the answer that best describes you.

0	1	2	3	4
Not at all	A little	Somewhat	A lot	Almost always

CHOCI: Part 2

In this section, each of the questions asks you about thoughts, ideas, or pictures that keep coming into your mind even though you do not want them to do so. They may be unpleasant, silly, or embarrassing. For example, some children have the repeated thought that germs or dirt are harming them or other people, or that something unpleasant may happen to them or someone special to them. These are thoughts that keep coming back, over and over again, even though you do not want them.

Please answer each question by putting a circle around the number that best describes how much you agree with the statement, or how much you think it is true of you. Please answer each item, without spending too much time on any one item. There are no right or wrong answers.

Example:	**Not at all**	**Somewhat**	**A lot**
I often have the same upsetting thought about death over and over again.	**1**	**2**	**3**

How much do you agree with each of the following statements?	**Not at all**	**Somewhat**	**A lot**
1. I often have the same upsetting thought about an accident over and over again.	1	2	3
2. I am often very upset by sudden feelings that I want to harm myself.	1	2	3
3. I often have bad thoughts that make me feel like a terrible person.	1	2	3
4. I often have horrible thoughts about my family being hurt that upset me very much.	1	2	3
5. I am often very upset by a sudden feeling that I am going to harm my family.	1	2	3
6. I often have doubts about whether I've made the right decision.	1	2	3
7. I often have the same upsetting picture in my head about death over and over again.	1	2	3
8. I often have mean thoughts that I feel are terrible, over and over again.	1	2	3

How much do you agree with each of the following statements?	Not at all	Somewhat	A lot
9. I often have the same horrible picture in my head about an accident.	1	2\	3
10. I have upsetting sexual thoughts over and over again, even though I don't want them.	1	2	3
11. I often have horrible thoughts about going crazy.	1	2	3
12. I always think that something terrible is going to happen and it will be my fault.	1	2	3
13. I often think that my bad thoughts are as awful as actually doing the bad thing.	1	2	3

Please list the three most severe **thoughts** that you often have **and can't stop thinking about.** For example, thinking about hurting someone, or thinking bad things about God.

1)_____

2)_____

3)_____

How much time do you spend thinking about these things? Please circle the answer that best describes you.

0	**1**	**2**	**3**	**4**
None	Less than 1 hr. a day (occassionally)	1-3 hrs. a day (part of a morning or afternoon)	3-8 hrs. a day (about half the time you're awake)	More than 8 hrs. a day (almost all the time you're awake)

How much do these thoughts get in the way of school or doing things with friends? Please circle the answer that best describes you.

0	**1**	**2**	**3**	**4**
Not at all	A little	Somewhat	A lot	Extreme

How much do these thoughts bother or upset you? Please circle the answer that best describes you.

0	**1**	**2**	**3**	**4**
Not at all	A little	Somewhat	A lot	Extreme

How hard do you try to stop the thoughts or ignore them? Please circle the answer that best describes you.

0	**1**	**2**	**3**	**4**
I always try to resist	I try to resist most of the time	I make some effort to resist	Even though I want to, I don't try to resist	I don't resist at all

When you try to fight the thoughts, can you beat them? How much control do you have over the thoughts? Please circle the answer that best describes you.

0	**1**	**2**	**3**	**4**
Complete control	Much control	Moderate control	Little control	No control

How much have you been avoiding doing anything, going any place, or being with anyone because of your thoughts? Please circle the answer that best describes you.

0	**1**	**2**	**3**	**4**
Not at all	A little	Somewhat	A lot	Almost always

Appendix E. Beck Depression Inventory

Beck's Depression Inventory

This depression inventory can be self-scored. The scoring scale is at the end of the questionnaire.

1.
- 0 I do not feel sad.
- 1 I feel sad
- 2 I am sad all the time and I can't snap out of it.
- 3 I am so sad and unhappy that I can't stand it.

2.
- 0 I am not particularly discouraged about the future.
- 1 I feel discouraged about the future.
- 2 I feel I have nothing to look forward to.
- 3 I feel the future is hopeless and that things cannot improve.

3.
- 0 I do not feel like a failure.
- 1 I feel I have failed more than the average person.
- 2 As I look back on my life, all I can see is a lot of failures.
- 3 I feel I am a complete failure as a person.

4.
- 0 I get as much satisfaction out of things as I used to.
- 1 I don't enjoy things the way I used to.
- 2 I don't get real satisfaction out of anything anymore.
- 3 I am dissatisfied or bored with everything.

5.
- 0 I don't feel particularly guilty
- 1 I feel guilty a good part of the time.
- 2 I feel quite guilty most of the time.
- 3 I feel guilty all of the time.

6.
- 0 I don't feel I am being punished.
- 1 I feel I may be punished.
- 2 I expect to be punished.
- 3 I feel I am being punished.

7.
- 0 I don't feel disappointed in myself.
- 1 I am disappointed in myself.
- 2 I am disgusted with myself.
- 3 I hate myself.

8.
- 0 I don't feel I am any worse than anybody else.
- 1 I am critical of myself for my weaknesses or mistakes.
- 2 I blame myself all the time for my faults.
- 3 I blame myself for everything bad that happens.

9.
- 0 I don't have any thoughts of killing myself.
- 1 I have thoughts of killing myself, but I would not carry them out.
- 2 I would like to kill myself.
- 3 I would kill myself if I had the chance.

10.
- 0 I don't cry any more than usual.
- 1 I cry more now than I used to.
- 2 I cry all the time now.
- 3 I used to be able to cry, but now I can't cry even though I want to.

11.

0 I am no more irritated by things than I ever was.
1 I am slightly more irritated now than usual.
2 I am quite annoyed or irritated a good deal of the time.
3 I feel irritated all the time.

12.

0 I have not lost interest in other people.
1 I am less interested in other people than I used to be.
2 I have lost most of my interest in other people.
3 I have lost all of my interest in other people.

13.

0 I make decisions about as well as I ever could.
1 I put off making decisions more than I used to.
2 I have greater difficulty in making decisions more than I used to.
3 I can't make decisions at all anymore.

14.

0 I don't feel that I look any worse than I used to.
1 I am worried that I am looking old or unattractive.
2 I feel there are permanent changes in my appearance that make me look unattractive
3 I believe that I look ugly.

15.

0 I can work about as well as before.
1 It takes an extra effort to get started at doing something.
2 I have to push myself very hard to do anything.
3 I can't do any work at all.

16.

0 I can sleep as well as usual.
1 I don't sleep as well as I used to.
2 I wake up 1-2 hours earlier than usual and find it hard to get back to sleep.
3 I wake up several hours earlier than I used to and cannot get back to sleep.

17.

0 I don't get more tired than usual.
1 I get tired more easily than I used to.
2 I get tired from doing almost anything.
3 I am too tired to do anything.

18.

0 My appetite is no worse than usual.
1 My appetite is not as good as it used to be.
2 My appetite is much worse now.
3 I have no appetite at all anymore.

19.

0 I haven't lost much weight, if any, lately.
1 I have lost more than five pounds.
2 I have lost more than ten pounds.
3 I have lost more than fifteen pounds.

20.

0	I am no more worried about my health than usual.
1	I am worried about physical problems like aches, pains, upset stomach, or constipation.
2	I am very worried about physical problems and it's hard to think of much else.
3	I am so worried about my physical problems that I cannot think of anything else.

21.

0	I have not noticed any recent change in my interest in sex.
1	I am less interested in sex than I used to be.
2	I have almost no interest in sex.
3	I have lost interest in sex completely.

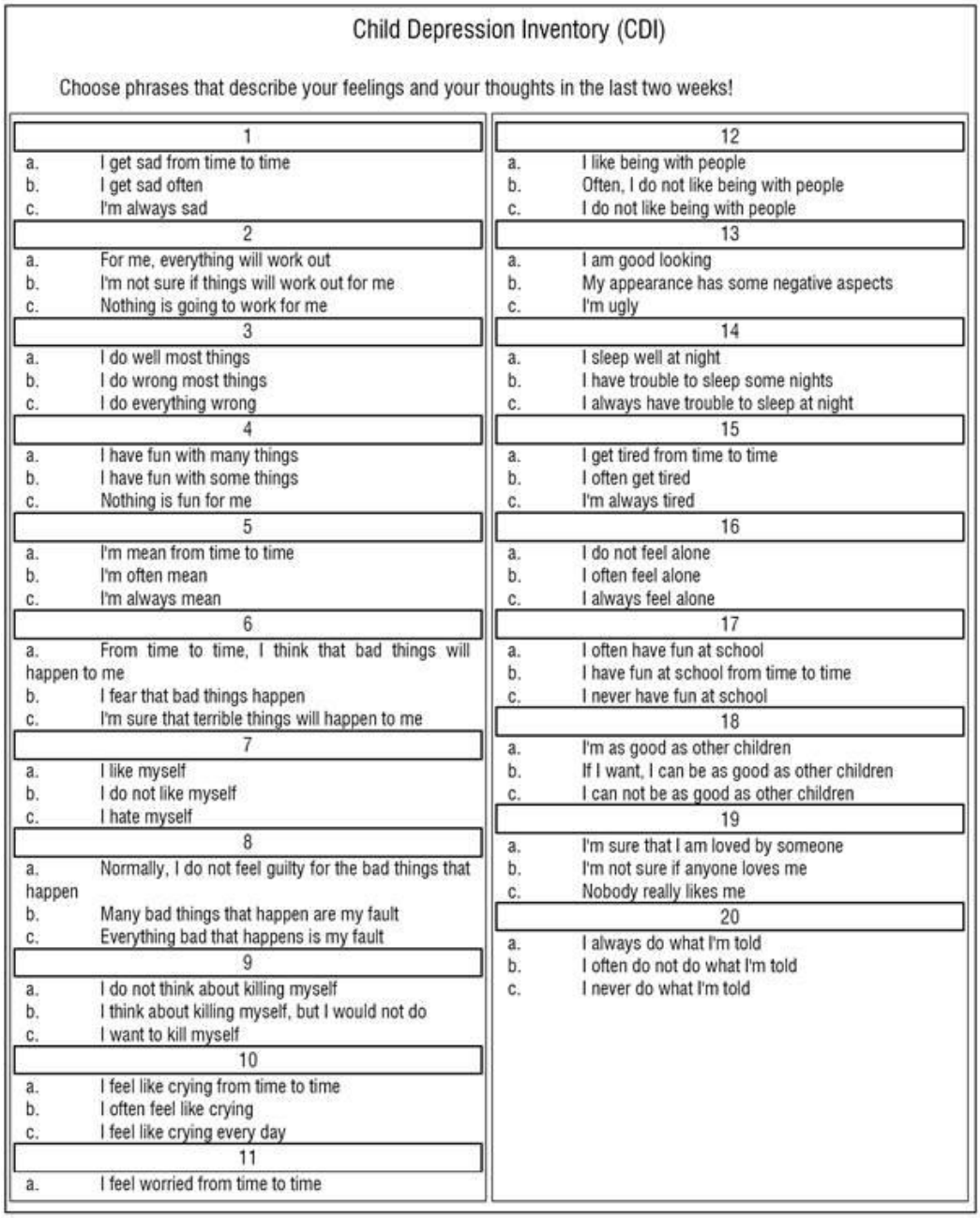
INTERPRETING THE BECK DEPRESSION INVENTORY

Now that you have completed the questionnaire, add up the score for each of the twenty-one questions by counting the number to the right of each question you marked. The highest possible total for the whole test would be sixty-three. This would mean you circled number three on all twenty-one questions. Since the lowest possible score for each question is zero, the lowest possible score for the test would be zero. This would mean you circles zero on each question. You can evaluate your depression according to the Table below.

Total Score	Levels of Depression
1-10	These ups and downs are considered normal
11-16	Mild mood disturbance
17-20	Borderline clinical depression
21-30	Moderate depression
31-40	Severe depression
over 40	Extreme depression

http://www.med.navy.mil/sites/NMCP2/PatientServices/SleepClinicLab/Documents/Beck_Depression_Inventory.pdf

Appendix E. Child Depression Inventory



Child Depression Inventory (CDI)

Choose phrases that describe your feelings and your thoughts in the last two weeks!

1	
a.	I get sad from time to time
b.	I get sad often
c.	I'm always sad
2	
a.	For me, everything will work out
b.	I'm not sure if things will work out for me
c.	Nothing is going to work for me
3	
a.	I do well most things
b.	I do wrong most things
c.	I do everything wrong
4	
a.	I have fun with many things
b.	I have fun with some things
c.	Nothing is fun for me
5	
a.	I'm mean from time to time
b.	I'm often mean
c.	I'm always mean
6	
a.	From time to time, I think that bad things will happen to me
b.	I fear that bad things happen
c.	I'm sure that terrible things will happen to me
7	
a.	I like myself
b.	I do not like myself
c.	I hate myself
8	
a.	Normally, I do not feel guilty for the bad things that happen
b.	Many bad things that happen are my fault
c.	Everything bad that happens is my fault
9	
a.	I do not think about killing myself
b.	I think about killing myself, but I would not do
c.	I want to kill myself
10	
a.	I feel like crying from time to time
b.	I often feel like crying
c.	I feel like crying every day
11	
a.	I feel worried from time to time

12	
a.	I like being with people
b.	Often, I do not like being with people
c.	I do not like being with people
13	
a.	I am good looking
b.	My appearance has some negative aspects
c.	I'm ugly
14	
a.	I sleep well at night
b.	I have trouble to sleep some nights
c.	I always have trouble to sleep at night
15	
a.	I get tired from time to time
b.	I often get tired
c.	I'm always tired
16	
a.	I do not feel alone
b.	I often feel alone
c.	I always feel alone
17	
a.	I often have fun at school
b.	I have fun at school from time to time
c.	I never have fun at school
18	
a.	I'm as good as other children
b.	If I want, I can be as good as other children
c.	I can not be as good as other children
19	
a.	I'm sure that I am loved by someone
b.	I'm not sure if anyone loves me
c.	Nobody really likes me
20	
a.	I always do what I'm told
b.	I often do not do what I'm told
c.	I never do what I'm told

Appendix F. Brief Conceptual Description of Latent Profile Analysis

The process of LPA involves specifying *a priori* how many latent profiles should be sought within the data. The latent variable upon which the profiles are based is calculated from the observed variables within the data; it is the product of these variables and an arbitrarily specified weight. The probability for each case of belonging to each latent profile is then calculated. The distribution of each group within the latent variable is assumed to be normal. The posterior probabilities can then be used to compare the distribution of the latent classes, based on the current weights, with the assumed normal distribution. After each iteration, the weights are adjusted to reduce the error between the assumed and observed distributions in the latent variable. The process continues until the distribution of the latent classes converges with the normal distribution (Berlin et al., 2014).

Appendix G. Latent Profile Analysis Code

```
Title: AN5prof

  Data: FILE IS C:\Users\ad626\Downloads\rLPA_AN.dat;

  VARIABLE: NAMES ARE iq ir dr rt cci vc1 vc2 vc3 vc4 cw3 cw4 ttach ttc4
bix_rev;

  MISSING ARE ALL (999);

  ! Seek c-classes solution

  CLASSES = c (5);

  ANALYSIS: TYPE = MIXTURE;

  ! LRTBOOT indicates number of draws (100) when bootstrapping p-value with
TECH14

  LRTBOOT=100;

  ! LRTSTARTS indicates the number of start values used in the initial stage and

  ! optimisations in the final stage when estimating the k-1 and k classes.

  ! Here, in the k-1 model, 2 random sets of start values are used in the initial

  ! stage and one optimisation is used in the final stage.

  ! In the k model, 100 random sets of start values are used in the initial stage

  ! and 500 optimisations should be used in the final stage.

  ! default starts option

  ! STARTS = 20 4;

  ! STITERATIONS = 10;

  ! enhance starts option for testing local maxima

  STARTS = 100 20;

  STITERATIONS = 40;

  LRTSTARTS = 2 1 100 50;

  ! use 4 processor threads

  PROCESSORS=4;

  MODEL:

  %OVERALL%
```

```
  iq;
  ir;
  dr;
  rt;
  cci;
  vc1;
  vc2;
  vc3;
  vc4;
  cw3;
  cw4;
  ttach;
  ttc4;
  bix_rev;
  SAVEDATA: SAVE=CPROBABILITIES;
  FILE IS AN5Ca.txt;
  ! OUTPUT: SAMPSTAT TECH1 TECH10 TECH11 TECH14;
  ! TECH 1 = request the arrays containing parameter
  !         specifications and starting values for all free parameters in the model.
  ! TECH 8 = request that the optimization history in estimating the model be
printed
  !         in the output.
  ! TECH 10 = request univariate, bivariate, and response pattern model fit
information
  !          for the categorical dependent variables in the model. This includes
  !          observed and estimated (expected) frequencies and standardized
residuals.
  ! TECH 11 = The TECH11 option is used in conjunction with TYPE=MIXTURE to
request
  !          the Lo-Mendell-Rubin likelihood ratio test of model fit
```

```
!       (Lo, Mendell, & Rubin, 2001) that compares the estimated model with a
!       model with one less class than the estimated model.
! TECH 14 = request a parametric bootstrapped likelihood ratio test (McLachlan &
!       Peel, 2000) that compares the estimated model to a model with one less
!       class than the estimated model.
OUTPUT: SAMPSTAT TECH14;
PLOT:
TYPE IS PLOT3;
SERIES = iq ir dr rt cci vc1 vc2 vc3 vc4 cw3 cw4 ttach ttc4 bix_rev (*);
```

Appendix H. Brief Conceptual Description of Neural Networks

A perceptron (Rosenblatt, 1957) is the simplest form of NN (Figure 4). Its inputs are the variables the network uses to predict the output. Independent weights (values between 0 and 1) are applied to each of the inputs and summed by the 'neuron' in the hidden layer. An activation function is then applied to define the range of possible output values.

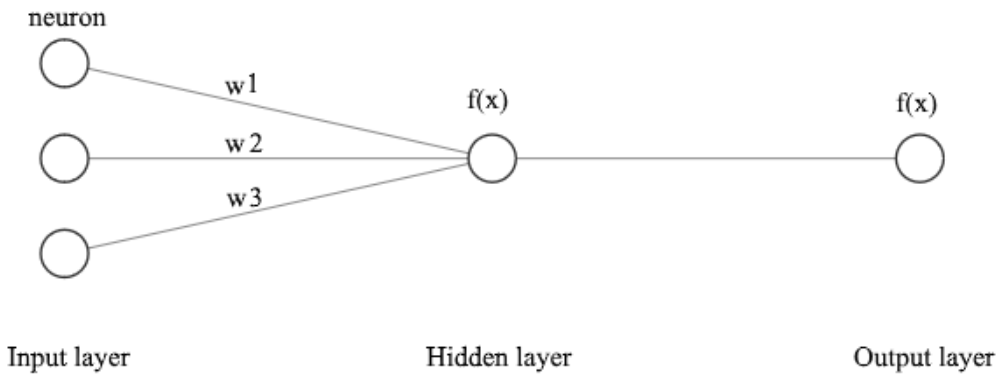


Figure 5. A Perceptron. w = weight. $f(x)$ is the activation function applied to the output of each hidden layer and output layer neuron. Different activation functions can be used at each layer of the model.

A modern NN builds upon the perceptron by using many neurons in the hidden layer and multiple hidden layers. Multi-layer networks can outperform single-layer networks but take much longer to train. Training is the iterative process of adjusting the weights to accurately predict an output from the inputs.

Appendix I. Neural Network Code

```
# Importing the libraries


import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

from scipy import stats


# Importing the dataset. Five datasets used to test different data models

# dataset = pd.read_csv('H1 ravello test data MINUS DX BMI EDE DEP OCD ANX AGE.csv')

dataset = pd.read_csv('H2 ravello test data MINUS DX BMI EDE DEP OCD ANX.csv')

# dataset = pd.read_csv('H3 ravello test data MINUS DX BMI EDE DEP OCD.csv')

# dataset = pd.read_csv('H4 ravello test data MINUS DX BMI EDE DEP.csv')

# dataset = pd.read_csv('H5 ravello test data MINUS DX BMI EDE.csv')

# dataset = pd.read_csv('H6 ravello test data MINUS DX BMI.csv')

# dataset = pd.read_csv('H7 ravello test data MINUS DX.csv')


# select from dataset which variables to include

# In python, the upperbound of a range is excluded.

X = dataset.iloc[:, 1:dataset.shape[1]].values

y = dataset.iloc[:, 0].values


# define number of features in set

features = X.shape[1]


# Splitting the dataset into the Training set and Test set
```

```
from sklearn.model_selection import train_test_split
# CHANGE RANDOM STATE TO 1 WHEN DOING REAL TESTING
X_train, X_test, y_train, y_test = train_test_split(X, y,
                                    test_size = 0.2,
                                    random_state = 1)

#####################################################
# Tuning the ANN
# Dropout regularisation to reduce overfitting if needed
# (apprent when large difference between difference in accuracy of
# training set and cv set)

from keras.wrappers.scikit_learn import KerasClassifier
from sklearn.model_selection import GridSearchCV
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Dropout
from keras.layers import LeakyReLU

# build the classifier NN as per the architecture above

def build_classifier(optimizer, neurons, divide):
    classifier = Sequential()
    classifier.add(Dense(units = neurons,
                         kernel_initializer = 'uniform',
                         activation = 'linear',
                         input_dim = features))
```

```
    # classifier.add(LeakyReLU(alpha=0.01))
    classifier.add(Dropout(rate = 0.1))
    classifier.add(Dense(units = round(neurons/divide),
                         kernel_initializer = 'uniform',
                         activation = 'linear'))
    # classifier.add(LeakyReLU(alpha=0.01))
    classifier.add(Dropout(rate = 0.1))
    classifier.add(Dense(units = 1, kernel_initializer = 'uniform',
                         activation = 'sigmoid'))
    classifier.compile(optimizer = optimizer,
                       loss = 'binary_crossentropy',
                       metrics = ['accuracy'])
    return classifier #output the classifier

# build classifier NN using parameters from build_classifier
# classifier parameters removed so GridSearch can find them computationally
classifier = KerasClassifier(build_fn = build_classifier)
# parameters is dictionary of parameters and values
# batch_size to test (25 and 32) based on

"""
parameters = {'batch_size': [32, 25, 10, 5, 2, 1],
              'epochs': [1000, 500, 250, 100],
              'optimizer': ['adam'],
              'neurons':[27, 14, 9, 7],
              'divide':[1, 2, 3, 4, 5, 6, 7, 8]}
# sans neurons took 48 hours to run
```

```
"""

parameters = {'batch_size': [5],
              'epochs': [500],
              'optimizer': ['adam'],
              'neurons':[27],
              'divide':[4]}

grid_search = GridSearchCV(estimator = classifier,
                           param_grid = parameters,
                           scoring = 'accuracy',
                           cv = 10,
                           n_jobs = 1)
grid_search = grid_search.fit(X_train, y_train)
best_parameters = grid_search.best_params_
best_accuracy = grid_search.best_score_

# change means and stds and params labels
# report parameter results
means = grid_search.cv_results_['mean_test_score']
stds = grid_search.cv_results_['std_test_score']
params = grid_search.cv_results_['params']
for mean, stdev, param in zip(means, stds, params):
    print("%f (%f) with: %r" % (mean, stdev, param))

#############################################################################
# Evaluating the ANN
```

```
import keras
from keras.wrappers.scikit_learn import KerasClassifier
from sklearn.model_selection import cross_val_score
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Dropout
from keras.layers import LeakyReLU

#build the classifier NN as per the architecture above
def build_linear_classifier(optimizer):
    classifier = Sequential()
    classifier.add(Dense(units = 27, kernel_initializer = 'uniform',
                         activation = 'linear',
                         input_dim = features))
    classifier.add(Dropout(rate = 0.1))
    classifier.add(Dense(units = 7, kernel_initializer = 'uniform',
                         activation = 'linear'))
    classifier.add(Dropout(rate = 0.1))
    classifier.add(Dense(units = 1, kernel_initializer = 'uniform',
                         activation = 'sigmoid'))
    classifier.compile(optimizer = optimizer,
                       loss = 'binary_crossentropy',
                       metrics = ['accuracy'])
    return classifier #output the classifier


def build_lrelu_classifier(optimizer):
```

```
    classifier = Sequential()
    classifier.add(Dense(units = 27, kernel_initializer = 'uniform',
                         activation = 'linear',
                         input_dim = features))
    classifier.add(LeakyReLU(alpha=0.01))
    classifier.add(Dropout(rate = 0.1))
    classifier.add(Dense(units = 7, kernel_initializer = 'uniform',
                         activation = 'linear'))
    classifier.add(LeakyReLU(alpha=0.01))
    classifier.add(Dropout(rate = 0.1))
    classifier.add(Dense(units = 1, kernel_initializer = 'uniform',
                         activation = 'sigmoid'))
    classifier.compile(optimizer = optimizer,
                       loss = 'binary_crossentropy',
                       metrics = ['accuracy'])
    return classifier #output the classifier

 # build classifier NN using parameters from build_classifier
linear_classifier = KerasClassifier(build_fn = build_linear_classifier,
                            batch_size = 5,
                            epochs = 500,
                            optimizer = 'adam')

lrelu_classifier = KerasClassifier(build_fn = build_lrelu_classifier,
                            batch_size = 5,
                            epochs = 500,
                            optimizer = 'adam')
```

```
# manual 5x2cv paired t test
# create arrays for storing linear model results
linear_accuracies = np.ones([5, 2])
linear_accuracies_output = np.ones([2, 1])
# create arrays for storing linear model results
lrelu_accuracies = np.ones([5, 2])
lrelu_accuracies_output = np.ones([2, 1])

# test models with 5x2-fold cross validation
for i in range(0,5):
    # shuffle training data
    X_train, X_test, y_train, y_test = train_test_split(X_train, y_train,
                                        test_size = 0,
                                        random_state = 1)
    # set cross-val
    linear_accuracies_output = cross_val_score(estimator = linear_classifier,
                        X = X_train,
                        y = y_train,
                        cv = 2,
                        n_jobs = 1)
    # update results array
    linear_accuracies[i] = np.transpose(linear_accuracies_output)

    lrelu_accuracies_output = cross_val_score(estimator = lrelu_classifier,
                        X = X_train,
                        y = y_train,
```

```
                    cv = 2,
                    n_jobs = 1)
    lrelu_accuracies[i] = np.transpose(lrelu_accuracies_output)
# t-test models with 5x2-fold cross validation
# error difference between models
error1 = linear_accuracies[:,0] - lrelu_accuracies[:,0]
error2 = linear_accuracies[:,1] - lrelu_accuracies[:,1]

mean_error = (error1 + error2)/2
mean_variance = ((error1 - mean_error)**2) + ((error2 - mean_error)**2)

# calculate t, means, and stds
mean_linear_accuracy = np.mean(linear_accuracies)
mean_lrelu_accuracy = np.mean(lrelu_accuracies)
mean_std_linear = np.std(linear_accuracies)
mean_std_lrelu = np.std(lrelu_accuracies)
t = error1[0]/(np.sqrt(sum(mean_variance)/5))
df = 5
p = 1 - stats.t.cdf(np.absolute(t),df=df)

# compare to critical t value for 5 df
if np.absolute(t) >= 2.015:
    print('significant')
else:
    print('non-significant')
```

Appendix J. The 5x2cvtest of Model Superiority

The 5x2cvtest compares the performance of two classification models (Dietterich, 1998). The test involves splitting the training dataset into two parts, $S_1$and S_2. Each algorithm, A and B is separately trained on datasets $S_1$and S_2, producing four error (*e*) scores, $e_A^{(1)}$, $e_B^{(1)}$, $e_A^{(2)}$, and $e_B^{(2)}$. The performance difference between algorithm A and B is calculated for each dataset:

$$e^{(1)} = e_A^{(1)} - e_B^{(1)} \tag{3}$$

$$e^{(2)} = e_A^{(2)} - e_B^{(2)} \tag{4}$$

The estimated variance between model performance is calculated with,

$$s^2 = \left(e^{(1)} - \bar{e}\right)^2 + \left(e^{(2)} - \bar{e}\right)^2 \tag{5}$$

Where $\bar{e}$ equals,

$$\bar{e} = \frac{\left(e^{(1)} + e^{(2)}\right)}{2} \tag{6}$$

The above is repeated five times using different random 50:50 splits of the data. The estimated *t*-value is then produced by

$$\tilde{t} = \frac{e_1^{(1)}}{\sqrt{\frac{1}{5}\sum_{i=1}^{5} s_i^2}} \tag{7}$$

Where $e_1^{(1)}$ is the error between the models taken from the first iteration, and s_i^2 is the estimated variance from the *i*th iteration. The absolute computed *t*-value is then compared against the critical value of *t* with five degrees of freedom.

Appendix K. Ethics Documentation

From: Øyvind RØ <oeyvro@ous-hf.no>

Subject: SV: permission to use the anonymised Ravello dataset for secondary analyses

To: 'IAN FRAMPTON' <ianframpton@mac.com>

To who it may concern:

I can confirm that in accordance with the original research ethics approval granted in Norway, secondary analyses of the fully anonymised Ravello research database can be conducted without seeking any further consent from the original participants.

Yours sincerely

Øyvind

Øyvind Rø, PhD, MD

Forskningsleder | Research Director

Regional seksjon spiseforstyrrelser | Regional Department for Eating Disorders

Klinikk psykisk helse og avhengighet | Division of Mental Health and Addiction

Oslo universitetssykehus HF | Helseregion Sør- Øst

ekspedisjon 23 01 62 30 | kontor 23 01 62 47

Professor i psykiatri / Professor Psychiatry

Institutt for klinisk medisin / Institute of Clinical Medicine

Universitetet i Oslo/ University of Oslo

Appendix L. Neural Network Parameter Testing

The parameters that produced the highest accuracy were batch size 5 and epochs 500 (Table 10). The greatest performance was obtained with a two-layer network with 27 neurons (Table 11). The optimal number of neurons in the second hidden layer was 7 (Table 12). See Figure X for a schematic of the final network design.

Table 10.

Epoch and Bach Parameter Testing

	Epochs			
Batch #	100	250	500	1000
1	0.695 (0.095)	0.704 (0.064)	0.695 (0.084)	0.707 (0.063)
2	0.698 (0.092)	0.719 (0.078)	0.710 (0.066)	0.716 (0.080)
5	0.707 (0.072)	0.707 (0.095)	**0.737 (0.087)**	0.725 (0.097)
10	0.719 (0.060)	0.683 (0.066)	0.701 (0.031)	0.701 (0.092)
25	0.713 (0.065)	0.725 (0.059)	0.734 (0.081)	0.719 (0.075)
32	0.716 (0.074)	0.734 (0.074)	0.707 (0.059)	0.734 (0.036)

Note. Batch # = number of training cases entered before weights adjusted. Epochs = number of times training data inputted into the network. Scores are mean accuracy from 10-folk cross-validation (0-1). Standard deviations in brackets. Greatest performance in bold.

Table 11.

Layer and Neuron Architecture Testing

	Layers			
Neurons	1	2	3	4
7	0.740 (0.109)	0.737 (0.070)	0.713 (0.064)	0.731 (0.077)
9	0.725 (0.091)	0.692 (0.074)	0.713 (0.081)	0.716 (0.067)
14	0.692 (0.073)	0.710 (0.071)	0.737 (0.068)	0.722 (0.090)
27	0.734 (0.084)	**0.746 (0.082)**	0.734 (0.048)	0.734 (0.091)

Note. ReLU = Rectifier linear unit. Scores are mean accuracy from 10-folk cross-validation (0-1). Standard deviations in brackets. Greatest performance in bold.

Table 12.

Second Hidden Layer Neuron Testing

	Second Layer Neurons							
	27	14	9	7	6	5	4	3
Accuracy	0.751	0.713	0.751	**0.769**	0.734	0.722	0.737	0.740
SD	0.057	0.074	0.072	**0.083**	0.087	0.074	0.073	0.084

Note. Scores are mean accuracy from 10-folk cross-validation (0-1). Standard deviations in brackets.

The above parameter testing was conducted with splitting randomisation disabled, meaning that the training data is split in a non-random manner. The performance of models during parameter specification is not comparable with the final model comparison, for which randomisation was enabled, as performance is affected by the split of the training sample.

Appendix M. AN and HC Demographic and Clinical Characteristics

Table 13.

Demographic and Clinical Characteristics Between AN Patients and HCs

	AN (*n* = 253)		HC (*n* = 170)		Test statistics		
	M	*SD*	*M*	*SD*	*F*	*p*	*d*
Age	15.62	1.82	14.52	2.26	32.01	< 0.001	0.25
BMI	16.04	1.87	20.85	3.59	254.39	< 0.001	0.78
Eating Restraint	3.22	1.88	0.55	0.88	326.73	< 0.001	0.60
Eating Concern	2.86	1.78	0.35	0.65	377.46	< 0.001	0.64
Weight Concern	3.48	1.97	0.90	1.16	251.53	< 0.001	0.55
Shape Concern	3.81	1.91	1.02	1.17	310.78	< 0.001	0.59
EDE Global	3.33	1.69	0.71	0.88	382.64	< 0.001	0.62
Depression	1.82	1.38	−0.55	0.79	554.00	< 0.001	0.70
Anxiety	1.28	1.17	−0.53	0.71	434.29	< 0.001	0.51
Obsessionality	1.84	1.47	−0.01	0.97	266.03	< 0.001	0.57

Note. AN = anorexia nervosa; HC = healthy control, BMI = body mass index, EDE = eating disorder examination. Brown-Forsythe ANOVA used. Cohen's d sizes: small 0.2, medium 0.5 and large 0.8 (Cohen, 1988).

Appendix N. AN and HC Neuropsychological Performance

Table 14.

Standardized Neuropsychological Performance Between AN Patients and HCs

	AN (*n* = 253)		HC (*n* = 170)		Test statistic		
	M	*SD*	*M*	*SD*	*F*	*p*	*d*
IQ	107.34	12.57	107.24	10.70	0.08	0.93	0.000
IR	−0.65	1.11	−0.58	1.18	0.461	.498	0.001
DR	−0.69	1.16	−0.53	1.19	1.942	.164	0.005
RT	−0.46	1.10	−0.20	1.23	5.235	.023	0.012
CCI	0.43	0.89	−0.02	1.01	23.312	.000	0.052
VC1	0.92	1.16	0.07	0.96	63.353	.000	0.131
VC2	1.16	1.02	0.89	1.08	6.759	.010	0.016
VC3	1.23	1.13	0.80	1.01	16.525	.000	0.038
VC4	1.23	1.07	0.89	0.90	11.327	.001	0.026
CW3	0.22	1.11	0.40	0.57	3.719	.054	0.003
CW4	0.38	0.90	0.48	0.73	1.473	.226	0.009
TT	0.02	0.70	0.19	0.78	5.007	.026	0.012
TMT	−0.15	0.95	0.14	0.80	11.160	.001	0.026
BSAT	−0.25	1.15	−0.05	0.86	3.783	.052	0.009

Note. AN = anorexia nervosa; HC = healthy control, IR = immediate recall, DR = delayed recall, RT = recognition, CCI = central coherence index, VC1-4 = Verbal Fluency conditions 1-4, CW3-4 = Color-Word Interference conditions 3 and 4, TT = Tower of London test, TMT = Trails Making Task, BSAT = Brixton Spatial Anticipation Test. All values are z-scores. Cohen's d sizes: small 0.2, medium 0.5 and large 0.8 (Cohen, 1988).

Appendix O. Neuropsychological Correlational Matrix

Table 15.

AN Neuropsychological Variables Correlational Matrix

		Statistic		Variable														
#	Var	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	AN	0.60	0.49	1														
2	IQ	107.30	11.84	0.00	1													
3	IR	-0.62	1.14	-0.03	0.36†	1												
4	DR	-0.63	1.17	-0.07	0.34†	0.89†	1											
5	RT	-0.36	1.16	-.11*	0.19†	0.33†	0.34†	1										
6	CCI	0.25	0.96	.23†	.13*	.27†	.27†	-0.09	1									
7	VC1	0.58	1.16	.36†	.24†	.16†	.11*	-0.01	.19†	1								
8	VC2	1.05	1.05	.13†	.11*	.16†	.16†	0.02	0.09	.54†	1							
9	VC3	1.06	1.10	.19†	.23†	.16†	.16†	0.05	.11*	.49†	.58†	1						
10	VC4	1.09	1.02	.16†	.21†	.14†	.15†	0.02	0.06	.44†	.55†	.91†	1					
11	CW3	0.42	0.84	-0.06	.20†	.10*	0.07	0.01	0.07	.20†	.19†	.26†	.28†	1				
12	CW4	0.29	0.93	-0.09	.14†	.15†	.14†	0.00	0.10	0.07	.12*	.18†	.20†	.48†	1			
13	TT	0.09	0.74	-.11*	0.07	.12*	.13†	0.02	0.08	-0.02	-0.01	0.05	0.03	.17†	.10*	1		
14	TC4	-0.03	0.90	-.16†	.24†	0.07	0.08	0.07	-0.00	.11*	.13†	.17†	.26†	.39†	.32†	.22†	1	
15	BRIX	-0.17	1.05	-0.09	0.09	.13*	.13†	.10*	0.04	0.01	0.01	0.05	0.04	-0.03	0.03	0.04	0.00	1

Note. AN = anorexia nervosa, var = variable, * = significant at 0.05 level (two-tailed), † = significant at 0.01 level (two-tailed). Correlations not corrected for multiple comparisons. AN = anorexia nervosa, IQ = IQ, IR = immediate recall, DR = delayed recall, RT = recognition, CCI = Central Coherence Index, VC1-4 = Verbal Fluency Conditions 1-4, CW3-4 = Color-Word Interference conditions 3 and 4, TT = Tower Test, TMT = Trail Making Task.

Appendix P. Dissemination Statement

The study results were disseminated locally at the Exeter Eating Disorder Research Group, chaired by Dr Ian Frampton. The empirical paper will be submitted for publication in the International Journal of Eating Disorders. For brevity, the submission will comprise the latent profile analysis only. Finally, study outcomes will be submitted for presentation at professional conferences.

Appendix Q. International Journal of Eating Disorders Submission Guidelines

1. SUBMISSION

Authors should kindly note that submission implies that the content has not been published or submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium. If there is a related paper under consideration at another journal, a copy of that paper should be submitted with the primary manuscript as supporting information.

Authors should follow the guidelines carefully; failure to do so will delay the processing of the manuscript. **Once the submission has been prepared in accordance with the Author Guidelines, manuscripts should be submitted online at mc.manuscriptcentral.com/ijed**. Authors unfamiliar with ScholarOne can find details on how to use the system here: www.wileyauthors.com/scholarone.

The submission system will prompt the author to use an ORCID iD (a unique author identifier) to help distinguish their work from that of other researchers. Details can be found elsewhere in these guidelines.

By submitting a manuscript to or reviewing for this publication, an individual's name, email address, and affiliation, and other contact details the publication might require, will be used for the regular operations of the publication, including, when necessary, sharing with the publisher (Wiley) and partners for production and publication. The publication and the publisher recognize the importance of protecting the personal information collected from users in the operation of these services, and have practices in place to ensure that steps are taken to maintain the security, integrity, and privacy of the personal data collected and processed. You can learn more at authorservices.wiley.com/statements/data-protection-policy.

2. AIMS AND SCOPE

The *International Journal of Eating Disorders*—A leading peer-reviewed journal in the fields of psychology, psychiatry, public health, and nutrition & dietetics.

Mission: With a mission to advance the scientific knowledge needed for understanding, treating, and preventing eating disorders, the *International Journal of Eating Disorders* publishes rigorously evaluated, high-quality contributions to an international readership of health professionals, clinicians, and scientists. The journal also draws the interest of patient groups and advocates focused on eating disorders, and many of the articles draw attention from mainstream media outlets.

Scope: Articles featured in the journal describe state-of-the-art scientific research on theory, methodology, etiology, clinical practice, and policy related to eating disorders, as well as contributions that facilitate scholarly critique and discussion of science and practice in the field. Theoretical and empirical work on obesity or healthy eating falls within the journal's scope inasmuch as it facilitates the advancement of efforts to describe and understand, prevent, or treat eating disorders. The *International Journal of Eating Disorders* welcomes submissions from all regions of the world and representing

all levels of inquiry (including basic science, clinical trials, implementation research, and dissemination studies), and across a full range of scientific methods, disciplines, and approaches.

3. MANUSCRIPT CATEGORIES AND REQUIREMENTS

The *International Journal of Eating Disorders* publishes the following contribution types:

1. Original Articles
2. Brief Reports
3. Clinical Case Reports
4. Reviews
5. An Idea Worth Researching
6. Commentaries

When uploading their manuscript, authors will be asked to complete a checklist indicating that they have followed the Author Guidelines pertaining to the appropriate article type. All word limits relate to the body of the text (i.e., not including abstract, references, tables and figures) and represent maximum lengths. Authors are encouraged to keep their manuscript as short as possible while communicating clearly.

1) Original Articles

These contributions report substantive research that is novel, definitive, or complex enough to require a longer communication. Only a subset of research papers is expected to warrant full-length format.

- Word Limit: 4,500 (excluding abstract, references, tables or figures)
- Abstract: 250 words.
- References: 60 are recommended; more are permissible, for cause.
- Figures/Tables: a maximum of 8 essential tables/figures, overall.

When preparing their manuscript, authors should follow the IMRaD guidelines (***I***ntroduction, ***M***ethods, ***R***esults, ***a***nd ***D***iscussion), which are recommended by the International Committee of Medical Journal Editors (ICMJE) (J. Pharmacol. Pharmacother. 2010, 1, 42–58). When preparing the Methods section, authors should refer to the Editorial Policy on Sample Size and Statistics.

4. PREPARING THE SUBMISSION

Parts of the Manuscript

The submission should be uploaded in separate files: 1) manuscript file; 2) figures; 3) Supporting Information file(s).

1. Manuscript File

The text file should contain all of the manuscript text, including the tables and figure legends. The text should be presented in the following order, with items i-v appearing on the Title Page:

1. Title

2. A short running title of less than 40 characters
3. The full names of all authors
4. The authors' institutional affiliations where the work was conducted, with a footnote for an author's present address if different to where the work was carried out
5. Word counts (abstract and main text, excl. tables and references)
6. Acknowledgements
7. Abstract and Keywords
8. Main text
9. References
10. Tables (each table complete with title and footnotes)
11. Figure legends

Title Page

On the title page, authors should list the title, the short running title, the full names of all authors, and their affiliations. Authors should also state the **number of words** contained in the abstract and the number of words of the manuscript (excluding tables and references).

Title

The title should be short and informative, containing major keywords related to the content. The title should not contain abbreviations (see Wiley's best practice SEO tips).

Authorship

For details on eligibility for author listing, please refer to the journal's Authorship policy outlined in Section 5 of these Author Guidelines.

Acknowledgments

Contributions from individuals who do not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.

Conflict of Interest Statement

Authors will be asked to provide a conflict of interest statement during the submission process. See the journal's policy on Conflict of Interest outlined in Section 5 of these Author Guidelines. Authors should ensure they liaise with all co-authors to confirm agreement with the final statement.

Abstract

The word maximum and abstract format varies by contribution type (see above). When an abstract is required, the abstract should be typed as a single paragraph. The journal requires **structured abstracts** with three exceptions: the journal will continue to use unstructured abstracts for Clinical Case Reports, Commentaries and "An Idea Worth Researching".

Structured abstracts should be organized as follows: **Objective**: briefly indicate the primary purpose of the article, or major question addressed in the study. **Method**: indicate the sources of data, give brief overview of methodology, or, if review article, how the literature was searched and articles selected for discussion. For research based articles, this section should briefly note study design, how participants were selected, and major study measures. **Results**: summarize the key findings. **Discussion**: indicate main clinical, theoretical, or research applications/implications.

Keywords

Please provide five to seven keywords. Keywords should be taken from those recommended by the US National Library of Medicine's Medical Subject Headings (MeSH) browser list at www.nlm.nih.gov/mesh.

Main Text

- Authors should refrain from using terms that are stigmatizing or terms that are ambiguous. For further explanation and examples, see the 2016 IJED article by Weissman et al. entitled "*Speaking of that: Terms to avoid or reconsider in the eating disorders field*" (DOI: 10.1002/eat.22528.)
- The text should be divided as outlined in Section 3 "Manuscript Categories and Requirements".
- Manuscripts reporting original research should follow the **IMRaD guidelines** (***I***ntroduction, (***M***ethods, ***R***esults, ***a***nd ***D***iscussion), which are recommended by the International Committee of Medical Journal Editors (ICMJE) (J. Pharmacol. Pharmacother. 2010, 1, 42–58).
- To facilitate evaluation by the Editors and Reviewers, each manuscript page should be numbered; the text should be double-spaced; and line numbers should be applied (restarting from 1 on each page). Instructions on how to implement this feature in Microsoft Word are given here.
- The journal uses US spelling. Authors may submit using any form of English as the spelling of accepted papers is converted to US English during the production process.
- Footnotes to the text are not allowed and any such material should be incorporated into the text as parenthetical matter.
- It is the primary responsibility of the authors to proofread thoroughly and ensure correct spelling and punctuation, completeness and accuracy of references, clarity of expression, thoughtful construction of sentences, and legible appearance prior to the manuscript's submission.
- Authors for whom English is not their first language are encouraged to seek assistance from a native or fluent English speaker to proof read the manuscript prior to submission. Wiley offers a paid service that provides expert help in English language editing—further details are given below.
- Articles reporting data taken from or deposited elsewhere should refer to the journal policy on Data Storage and Documentation in Section 5 (below).

References

References in all manuscripts should follow the style of the American Psychological Association (6th edition), except in regards to spelling. The APA website includes a range of resources for authors learning to write in APA style, including An overview of the Publication Manual of the American Psychological Association, Sixth Edition;

includes free tutorials on APA Style basics and an APA Style Blog. Please note APA referencing style requires that a Digital Object Identifier (DOI) be provided for all references where available.

Tables

Each table must be numbered in order of appearance in the text with Arabic numerals and be cited at an appropriate point in the text. Tables should be self-contained and complement, not duplicate, information contained in the text. They should be editable (i.e., created in Microsoft Word or similar), not pasted as images. Legends should be concise but comprehensive—the table, legend, and footnotes must be understandable without reference to the text. All abbreviations must be defined in footnotes. Footnote symbols: †, ‡, §, ¶, should be used (in that order) and *, **, *** should be reserved for P-values. Statistical measures such as standard deviation (SD) or standard error of the mean (SEM) should be identified in the headings. The journal's Editorial Policy on Sample Size and Statistics is given in Section 5.

Figure Legends/Captions

Each figure caption should have a brief title that describes the entire figure without citing specific panels, followed by a description of each panel. Captions should be concise but comprehensive—the figure and its caption must be understandable without reference to the text. Be sure to explain abbreviations in figures even if they have already been explained in-text. Axes for figures must be labeled with appropriate units of measurement and description. Include definitions of any symbols used and units of measurement.

2. Figures

Although authors are encouraged to send the highest quality figures possible, for peer-review purposes, a wide variety of formats, sizes, and resolutions are accepted. Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post-acceptance figure requirements.

Helvetica typeface is preferred for lettering within figures. All letters, numbers and symbols must be at least 2 mm in height. Courier typeface should be used for sequence figures. Figures should be numbered consecutively with Arabic numerals, and they should be numbered in the order in which they appear in the text.

Figures should be submitted as electronic images to fit either one (55 mm, 2 3/16", 13 picas), two (115 mm, 4 1/2", 27 picas), or three (175 mm, 6 7/8", 41 picas) columns. The length of an illustration cannot exceed 227 mm (9"). Journal quality reproduction requires grey scale and color files at resolutions of 300 dpi. Bitmapped line art should be submitted at resolutions of 600–1200 dpi.

Figures submitted in color will be reproduced in color online free of charge. Authors wishing to have figures printed in color in hard copies of the journal will be charged a fee by the Publisher; further details are given elsewhere in these Author Guidelines. Authors should note however, that it is preferable that line figures (e.g., graphs) are supplied in black and white so that they are legible if printed by a reader in black and white.

3. Supporting Information Files(s)

Supporting Information is information that is supplementary and not essential to the article, but provides greater depth and background. Examples of such information include more detailed descriptions of therapeutic protocols, results related to exploratory or post-hoc analyses, and elements otherwise not suitable for inclusion in the main article, such as video clips, large sections of tabular data, program code, or large graphical files. It is *not* appropriate to include, in the Supporting Information, text that would normally go into a discussion section; all discussion-related material should be presented in the main article.

Because the Supporting Information is separate from the paper and supplementary in nature, the main article should be able to be read as a stand-alone document by readers. Reference to the Supporting Information should be made in the text of the main article to provide context for the reader and highlight where and how the supplemental material contributes to the article.

Should authors wish to provide supplementary file(s) along with their article, these materials *must* be included upon submission to the journal. If such materials are added to the submission as a result of peer review, i.e., during a revision, then the authors should bring this to the attention of the editor in their response letter. If accepted for publication, Supporting Information is hosted online together with the article and appears without editing or typesetting.

Note: Authors are encouraged to utilize publicly available data repository for data, scripts, or other artefacts used to generate the analyses presented in the paper; in such cases, authors should include a reference to the location of the material within their paper.

General Style Points

The following points provide general advice on formatting and style.

- **Terminology**: Terms such as “anorexics” or “bulimics” as personal pronouns, referring to groups of individuals by their common diagnosis, should be avoided. Terms like “individuals with anorexia nervosa”, “people with bulimia nervosa”, or “participants with eating disorders” should be used instead. Note, “participants” should be used in place of “subjects”.
- **Abbreviations**: In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Initially, use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only.
- **Units of measurement**: Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website at www.bipm.fr for more information about SI units.
- **Numbers** under 10 should be spelt out, except for: measurements with a unit (8 mmol/L); age (6 weeks old), or lists with other numbers (11 dogs, 9 cats, 4 gerbils).
- **The word “data”** is plural; therefore, text should follow accordingly (for example, “The data show…the data are … the data were…”).
- **Sex/Gender & Age**: When referring to sex/gender, “males" and “females” should be used only in cases where the study samples include both children

(below age 18) and adults and only if word limit precludes using terms such as "male participants/female participants," "female patients/male patients"; when the participants comprise adults only, the terms "men" and "women" should be used. In articles that refer to children, "boys" and "girls" should be used.

- **Trade Names**: Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name and the name and location of the manufacturer in parentheses.
- **Statistics**: Authors should adhere to the journal's policy on Sample Size and Statistics when reporting studies. For information on how to present p values and other standard measurements see IJED Statistical Formatting Requirements.

5. EDITORIAL POLICIES AND ETHICAL CONSIDERATIONS

Editorial Policy on Sample Size and Statistics

The Methods section should include a statement about sample selection, response rate, and other factors that would impact selection or response bias and, in turn, representativeness of the sample. Inclusion of small samples requires justification and authors should be mindful of the recommendations concerning minimal sample sizes in subfields (e.g., genetic research, instrument development, etc., where adequate samples may number in the hundreds). Authors also are asked to provide information about reliability and validity of study measures as applicable to their sample.

Manuscripts reporting statistical tests without effect size estimates may be rejected without review.

Data Sharing and Data Accessibility

All accepted manuscripts are required to publish a data availability statement to confirm the presence or absence of shared data.

The *International Journal of Eating Disorders* recognizes the many benefits of archiving research data. *IJED* expects you to archive all the data from which your published results are derived in a public repository. The repository that you choose should offer you guaranteed preservation (see the registry of research data repositories at https://www.re3data.org/) and should help you make it findable, accessible, interoperable, and re-useable, according to FAIR Data Principles.

The *International Journal of Eating Disorders* notes that FAIR data sharing allows for access to shared data under restrictions (e.g., to protect confidential or proprietary information) but notes that the FAIR principles encourage you to share data in ways that are as open as possible (but that can be as closed as necessary).

If you have shared data, this statement will describe how the data can be accessed, and include a persistent identifier (e.g., a DOI for the data, or an accession number) from the repository where you shared the data. If you cannot share the data described in your manuscript, for example for legal or ethical reasons, or do not intend to share the data then you must provide the appropriate data availability statement. Sample statements are available here. If published, all statements will be placed in the heading of your manuscript.

Human Studies and Subjects

For manuscripts reporting studies that involve human participants, a statement identifying the ethics committee that approved the study and confirmation that the study conforms to recognized standards is required, for example: Declaration of Helsinki; US Federal Policy for the Protection of Human Subjects ; or European Medicines Agency Guidelines for Good Clinical Practice.

Every effort should be taken to ensure the anonymity of the patient concerned, and any clinicians not involved as authors. If there is any potentially identifiable information, then it is the responsibility of the authors to seek and obtain approval from the local Institutional Review Board (IRB) (or equivalent) for the case to be reported, and a copy of that approval should be made available to the Editor on request.

Images and information from individual participants will only be published where the authors have obtained the individual's free prior informed consent. Authors do not need to provide a copy of the consent form to the publisher; however, in signing the author license to publish, authors are required to confirm that consent has been obtained. Wiley has a standard patient consent form available for use.

Research Reporting Guidelines

Accurate and complete reporting enables readers to fully appraise research, replicate it, and use it. Authors are encouraged to adhere to any research reporting standards relevant to their study. A list of the most well-known guidelines is given here:

- Consolidated Standards of Reporting Trials (CONSORT)
- Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)
- Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)
- PRISMA Protocols (PRISMA-P)
- STrengthening the Reporting of OBservational studies in Epidemiology (STROBE)
- CARE: Guidelines to increase the accuracy, transparency, and usefulness of case reports
- Consolidated criteria for reporting qualitative research (COREQ) by Tong et al. (*Int. J. Qual. Health Care* (**2007**) *19*(6): 349–357)
- STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies
- TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
- Consolidated Health Economic Evaluation Reporting Standards (CHEERS) by Husereau et al. (*BMC Medicine*(**2013**) *11*: 80; DOI: 10.1186/1741-7015-11-80)
- The EQUATOR Network: an author's one-stop-shop for writing and publishing high-impact health research
- FORCE11: Recommended reporting guidelines for life science resources
- ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines
- Guidance for the Description of Animal Research in Scientific Publications from the US National Research Council's Institute for Laboratory Animal Research
- The Gold Standard Publication Checklist from Hooijmans et al. (ATLA (**2010**) *38*: 167–182)

Conflict of Interest

The journal requires that all authors disclose any potential sources of conflict of interest. Any interest or relationship, financial or otherwise that might be perceived as influencing an author's objectivity is considered a potential source of conflict of interest. These must be disclosed when directly relevant or directly related to the work that the authors describe in their manuscript.

Potential sources of conflict of interest include, but are not limited to: employment at a for-profit treatment center where data collection occurred, employment at a for-profit corporation if the corporation manufactures or sells products used in the research (e.g., medications; equipment used in a treatment tested as part of the research), patent or stock ownership, membership of a company board of directors, membership of an advisory board or committee for a company, and consultancy for or receipt of speaker's fees from a company.

The existence of a conflict of interest does not preclude publication. It is the responsibility of the corresponding author to review this policy with all authors and collectively to disclose with the submission ALL pertinent commercial and other relationships. These conflicts of interest should be disclosed in the relevant section of the submission questionnaire and in the manuscript. If the authors have no conflict(s) of interest to declare, they must also state this.

Funding

Authors should list all funding sources in the Acknowledgments section. Authors are responsible for the accuracy of their funder designation. If in doubt, please check the Open Funder Registry for the correct nomenclature: www.crossref.org/services/funder-registry.

Authorship

The list of authors should accurately illustrate who contributed to the work and how. All those listed as authors should qualify for authorship according to the following criteria:

1. Have made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data;
2. Been involved in drafting the manuscript or revising it critically for important intellectual content;
3. Given final approval of the version to be published. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content; and
4. Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section (for example, to recognize contributions from people who provided technical help, collation of data, writing assistance, acquisition of funding, or a department chairperson who provided general support). Prior to submitting the article all authors should agree on the order in which their names will be listed in the manuscript.

Joint first or senior authorship: In the case of joint first authorship, a footnote should be added to the author listing, e.g. ‘X and Y should be considered joint first author’ or ‘X and Y should be considered joint senior author.’

Authorship responsibilities: Eligibility for authorship requires that authors have made substantive contributions to the work described in the manuscript, have read and approved the manuscript in its current form, and have approved the ordering of authorship. All authors agree that, once a manuscript has been submitted, the subsequent addition, removal or change of authorship order requires the approval of all authors (including making such changes as part of a resubmission). Such changes in revised versions of a manuscript should be brought to the attention of the Editor in the response letter.