CHILD NEUROPSYCHOLOGY
2019, VOL. 25, NO. 6, 816–835
https://doi.org/10.1080/09297049.2018.1536200

Routledge
Taylor & Francis Group

Executive functioning in adolescent anorexia nervosa: Neuropsychology versus self- and parental-report

Laura Rebecca Herbrich, Viola Kappel, Sibylle Maria Winter* and Betteke Maria van Noort *

Department for Child and Adolescent Psychiatry, Psychosomatic Medicine and Psychotherapy, Charité – Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

ABSTRACT

There is limited research concerning the relationship between neuropsychological assessment and self-report of executive functioning in adolescent anorexia nervosa (AN); available studies demonstrate only low to moderate correlations. Therefore, this study examines the association between neuropsychological test performance and self-report in AN. Forty adolescent inpatients with AN completed an extensive neuropsychological assessment, including set-shifting, central coherence, and questionnaires assessing executive functioning in daily life (BRIEF-SR). Their parents filled out an analog version (BRIEF-PF). Statistical analyses revealed low to medium positive and negative correlations between neuropsychological measures and BRIEF subscales. Similarly, self- and parental ratings were only slightly positively correlated, with patients scoring significantly higher than their parents on two subscales. The results support previous findings of modest correlations between self-report and performance-based testing and emphasize the importance of a multiple format assessment of executive functioning in adolescent AN.

ARTICLE HISTORY
Received 29 June 2017
Accepted 4 October 2018

KEYWORDS
Anorexia nervosa; adolescents; neuropsychological performance; executive functioning; self-report

Anorexia nervosa (AN) is a severe mental disorder that is characterized by extensive weight loss, restrictive eating patterns, disturbances in body image, and widespread endocrine consequences (American Psychiatric Association, 2000). The illness is associated with serious medical complications, difficulties in psychological and emotional functioning and often tends to have a chronic course as patients typically remain ambivalent about treatment (Agras, Crow, Mitchell, Halmi, & Bryson, 2009, Arcelus, Haslam, Farrow, & Meyer, 2013, Arcelus, Mitchell, Wales, & Nielsen, 2011, Katzman, 2005). Its etiology is considered to be multifactorial and AN tends to set in during adolescence, a period of heightened vulnerability in terms of brain, behavioral, cognitive, and psychological development as well as endocrine changes (Blakemore, Burnett, & Dahl, 2010, Herpertz-Dahlmann, 2015). From a clinician's viewpoint, AN patients exhibit rigid and rule-bound behavioral patterns that manifest in a preoccupation with

CONTACT Laura Rebecca Herbrich laura.herbrich@charite.de Department for Child and Adolescent Psychiatry, Psychosomatic Medicine and Psychotherapy, Charité – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany
*These authors share senior authorship.

food and body as well as other noneating disorder behaviors such as order, planning, and organization. In terms of the possible underpinnings of this behavioral rigidity, an increasing amount of research has been put into the role of neuropsychological performance as a mediating factor between underlying neurobiology and psychological functioning (Kaye, Fudge, & Paulus, 2009, Kidd and Steinglass, 2012).

For adult AN patients, weaknesses in set-shifting and central coherence have been reported (Lang, Lopez, Stahl, Tchanturia, & Treasure, 2014, Wu et al., 2014). Cognitive set-shifting refers to the ability to shift between multiple tasks and change behavior in response to altering rules (Miyake et al., 2000). Weaknesses in central coherence describe the combination of a more detail-focused approach and poor global integration (Lang et al., 2014). These cognitive weaknesses tend to persist after weight gain and are likely to play a role in the vulnerability and maintenance of AN (Danner et al., 2012, Kanakam, Raoult, Collier, & Treasure, 2013, Shott et al., 2012, Talbot, Hay, Buckett, & Touyz, 2015, Tenconi et al., 2010). For adolescents with AN, the available data appears inconsistent and inefficiencies in cognitive functioning seem to be less pronounced than in adults with AN (Bentz et al., 2017, Kjaersdam Telleus et al., 2015). There are some findings similar to that of adult AN samples (Allen et al., 2012, Lang et al., 2015, McAnarney et al., 2011, Tchanturia et al., 2012), whereas other findings suggest neuropsychological performance comparable to that of healthy controls (Andres-Perpina et al., 2011, Fitzpatrick, Darcy, Colborn, Gudorf, & Lock, 2012, Kjaersdam Telleus et al., 2015, Lang, Stahl, Espie, Treasure, & Tchanturia, 2014, van Noort, Pfeiffer, Ehrlich, Lehmkuhl, & Kappel, 2016).

The assessment of neuropsychological functioning is generally carried out with standardized tests, most of which were originally developed for use with adult patients with brain damage. Likewise, inefficiencies in behavioral aspects of executive functioning are measured with neuropsychological tests or self-report questionnaires. However, it is questioned whether the available neuropsychological measures are sensitive enough to measure inflexible day-to-day behavior in other patient samples and if flexibility in everyday-life situations is generally quantifiable using standardized tests (Chaytor and Schmitter-Edgecombe, 2003). The well-structured nature of a neuropsychological assessment, in which the examiner provides guidance and cueing, may also relieve the participant of the need to fully exercise executive functioning as is done in real life, thus making it difficult to assess this competence solely by test performance (Gioia and Isquith, 2004). Furthermore, most tests assess accuracy and/or response time which results in global scores that do not allow the isolation of specific domains of executive functions (Jurado and Rosselli, 2007). In addition to methodological problems, the question of ecological validity is raised. Ecological validity refers to the degree of generalizability of an observed behavior in a standardized test environment to natural behavior in the real world (Schmuckler, 2001). In this regard, it is assumed that poor performances in test situations will lead to poor performances aside from the test environment. Yet, there is little evidence that supports this assumption and data obtained in a laboratory needs to be critically examined when attempting to predict behaviors outside of a test setting (Sbordone, 1996).

Given these restrictions and aiming at a more accurate clinical assessment, research has led to the development of relatively new measures, which complement existing neuropsychological tests and claim an increased level of ecological validity. The

"Behavior Rating Inventory for Executive Functioning" (BRIEF) (Drechsler and Steinhausen, 2013, Guy, Isquith, & Gioia, 2004) is one such example which provides a comprehensive assessment of children's and adolescents' performance of executive functioning in daily life situations. The questionnaire exists in a self- and parental-report as well as a teacher's version. The BRIEF has been used in several adolescent patient samples, such as attention deficit-hyperactivity disorder (ADHD) (Gioia, Isquith, Guy, & Kenworthy, 2000, Mahone et al., 2002, McCandless and O'Laughlin, 2007), brain injuries (Gioia and Isquith, 2004, Mahone, Zabel, Levey, Verda, & Kinsman, 2002, Wilson, Donders, & Nguyen, 2011) and psychosis (Niendam, Horwitz, Bearden, & Cannon, 2007). For AN patients, the data is limited, though recent studies evaluating cognitive remediation therapy (CRT) for adolescents have increasingly employed the BRIEF in addition to neuropsychological measures (Dahlgren, Lask, Landro, & Rø, 2014, Giombini, Moynihan, Turco, & Nesbitt, 2016, van Noort, Kraus, Pfeiffer, Lehmkuhl, & Kappel, 2015). CRT is a novel cognitive training program specifically tailored to address cognitive inefficiencies seen in AN (Davies and Tchanturia, 2005, Tchanturia, Lounes, & Holttum, 2014). McAnarney et al. (2011) found that both AN patients and their parents report more problems on behavioral and cognitive components of the BRIEF when compared to controls. Dahlgren et al. (2014) compared patient- and parental-report after receiving CRT and found both ratings to fall within the normal range before and after CRT, though with only moderate agreement between the reports. To date, there is only one study examining the relationship between self- and parental-report of executive functioning and test performance in adolescents with AN which reports low to moderate correlations (Stedal and Dahlgren, 2015). This same finding also applies for studies in brain injuries (Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002, Vriezen and Pigott, 2002) and ADHD or associated disorders (Mahone et al., 2002, McAuley, Chen, Goos, Schachar, & Crosbie, 2010, Toplak, Bucciarelli, Jain, & Tannock, 2009).

Following inconsistent findings of neuropsychological functioning in adolescent patients with AN and given the moderate correlation between everyday behavior as rated by the patients and test assessment, the current study aims to elucidate (1) the relationship between different neuropsychological measures as well as (2) the relationship between neuropsychological measures and patient self-report and (3) parental report in AN. Additionally, (4) the relationship between patient- and parental-report findings of executive functioning is explored.

Method

Study design and participants

By using a cross-sectional design, the data collection took place between June 2014 and December 2016. All participants were patients seeking inpatient treatment at our specialized unit for eating disorders at the Department for Child and Adolescent Psychiatry, Psychosomatic Medicine and Psychotherapy of the Charité-Universitätsmedizin Berlin. Inclusion criteria for participation were defined as (1) a current typical AN as outlined in the 10th edition of the International Classification of Diseases and Related Health Problems (ICD-10) (Remschmidt, Schmidt, & Poustka, 2006); (2) an age between 11

and 17.11 years; (3) female gender; and (4) inpatient treatment. Exclusion criteria were a history of or current substance abuse or dependence and an intelligence quotient (IQ) below 85. In light of the female-to-male ratio in clinical samples (1:10–12), the current article focusses solely on the female gender (Currin, Schmidt, Treasure, & Jick, 2005, Hoek and van Hoeken, 2003). However, it should be acknowledged that males also suffer from AN.

All female inpatients who were consecutively admitted to our specialized unit and fulfilled the in- and exclusion criteria were approached for participation within the first two weeks of treatment. Detailed information about the study and procedure was given to all patients and their legal guardians and, before participation, written informed consent was obtained. Ethical approval for this study was obtained from the ethical committee of the Charité-Universitätsmedizin Berlin, Campus Virchow Klinikum (EA2/026/11). The patients received no financial compensation.

Assessment

The patients completed an extensive assessment battery, including structured interviews, self-report questionnaires and neuropsychological tests, which are described in the assessment sub-sections below. As comorbidity is common among adolescents with AN (Bühren et al., 2014, Godart et al., 2007, Jaite, Hoffmann, Glaeske, & Bachmann, 2013), all other general psychiatric axis I-disorders apart from the eating disorder were taken from the patients' medical records, which were based on clinical consensus of the eating disorders specialized team supervised by the senior child and adolescent psychiatrist of the clinic. The body mass index (BMI) was calculated using standardized techniques (kg/m^2). As the 10^{th} BMI-percentile marks the critical body weight threshold for children and adolescents (Holtkamp and Herpertz-Dahlmann, 2005), BMI-percentiles were also derived (Kromeyer-Hauschild et al., 2001). The assessment was conducted by two of the authors (LH, BvN) or by trained master's students under supervision and took place during the first three weeks of treatment in an acute state of illness. The assessment was standardized and conducted individually on two consecutive days (one hour each). The parental questionnaires were filled out during the first three weeks of treatment. To provide information regarding the socioeconomic status (SES), parental education and their current occupation was used to calculate Hollingshead four factor index of social status (Hollingshead, 1975). Scores range from 8 to 66 indicating a low (<15), middle (16–60), or high (>61) status. Information on the SES allows for a better interpretation and comparability with other studies. Moreover, SES has been associated with cognitive abilities both directly and indirectly, for example, developmental opportunities, verbal skills, nutrition (Lezak, Howieson, Bigler, & Tranel, 2012).

Eating disorder assessment

A current AN diagnosis was confirmed via the 'Structured Inventory for Anorexic and Bulimic Eating Disorders for DSM-IV and ICD-10′ (SIAB-EX; Fichter and Quadflieg, 1999). As is common in Germany, the ICD-10 diagnoses for AN (F50.00, F50.01) were used in the current study. The SIAB-EX assesses different eating disorder diagnoses and the severity of symptoms for the time points ≪ current ≫ (last three months) and ≪ past ≫

(time up to three months before the interview). High inter-rater reliability (r = .81 (current) and r = .85 (past)) as well as high internal consistency (α = .93) are reported (Fichter and Quadflieg, 2001).

Eating disorder psychopathology was measured using the German version of the 'Eating Disorder Inventory-2' (EDI-2; Paul and Thiel, 2004). The 91-item questionnaire uses a 6-point Likert scale and includes 11 subscales and a global score. For the current study, the global mean score served as an indicator for the severity of eating disorder symptomatology, with higher scores indicating more psychopathology. Internal consistency lies between α = .73 and .93 (patient sample) and test-retest reliability between r = .81 and .89. Validity of the EDI-2, that is, content, convergent and discriminant validity, were all sufficient (Paul and Thiel, 2004).

Neuropsychological assessment

The selection of neuropsychological tests followed evidence from the latest systematic reviews (Lang et al., 2014, 2014, Wu et al., 2014) and each included two widely used measures for set-shifting and central coherence, which have been identified as sensitive in earlier studies with AN. First of all, IQ was assessed by the "Culture Fair Test 20 – revised" (CFT 20-R; Weiß, 2006), which is a non-verbal test assessing fluid intelligence. The CFT 20-R includes two analogous parts that each consist of four subtests (sequence completion, classifications, matrices and topology). The test shows high internal consistency (α between .93 and .95) and good test-retest reliability (r between .80 and .82). Moreover, its implementation time is short in order to keep the overall testing duration limited. More common IQ tests with a short duration time, that is, WASI-II, are not validated and available in German.

Cognitive set-shifting was measured using the 'Trail Making Test Condition 4′ (TMT-4) of the "Delis and Kaplan Executive Function System" (D-KEFS; Delis, Kaplan, & Kramer, 2001) and the "Wisconsin Card Sorting Test" (WCST;(Drühe-Wienholt and Wienholt, 2004, Heaton, Chelune, Kay, & Curtiss, 1993). The TMT-4 is administered using pen and paper and prompts patients to connect a sequence of numbers and letters in ascending order as quickly as possible (e.g., 1-A-2-B). The time taken to complete the task is the indicator for set-shifting. Internal consistency ranges between α = .57 and .81. Reliability (test–retest) was r = .38 over all age groups (8–80 years). The WCST is a computer-based test that consists of 96 different stimulus cards (German version), which the patients have to match according to changing rules. Patients are given feedback after each trial; however, the classification rule (color, shape, number) changes every ten cards and via implicit learning they must quickly adopt the new rule. Set-shifting is assessed by the number of perseverative errors (when patients keep applying the former classification rule). Internal consistency is good and reliability lies at r = .77 (Drühe-Wienholt and Wienholt, 2004).

Central coherence ability was assessed with the "Rey Complex Figures Test" (RCFT) (Meyers and Meyers, 1995) and the "Group Embedded Figures Test" (GEFT) (Witkin, Oltman, Raskin, & Karp, 1971). The RCFT consists of a complex figure composed of 18 global and local elements and requires the patients to copy the figure without a time limit. The instructor follows the patients' drawing process using a Flowchart-method. The scoring system follows Booth (2006) and evaluates the order in which the elements are drawn (global or detailed) and the style of the

drawing (coherent or fragmented). Scoring criteria include the "Order of Construction Index" (OCI, range 0–3.33), the "Style Index" (SI, range 0–2) and the overall score "Central Coherence Index" (CCI, range 0–2) that provide information concerning either a detail oriented or global approach of the task. A higher score on these three indices resembles more global processing strategies (Booth, 2006). The GEFT is administered using pen and paper and requires the patients to identify simple figures that are hidden within a progressively more complicated figure. The test assesses the ability to disembed certain information from surrounding irrelevant information. The number of correctly traced figures (range 0–18) serves as an outcome measure and higher scores indicate more field independence as opposed to a field dependent approach. The GEFT shows an internal consistency reliability index of α = .82 (Witkin et al., 1971).

Patient self-report of executive functioning

All patients filled out the German version of the "Behavior Rating Inventory for Executive Functioning – Self-Report" (BRIEF-SR) (Drechsler and Steinhausen, 2013, Guy et al., 2004), which is an 80-item questionnaire assessing executive functioning in everyday-life during the last six months, including daily-life situations at school and at home. The German version of the BRIEF-SR provides age and gender corrected norms for children and adolescents between 11 and 16.11 years. Although six patients were slightly older (maximum 17.8 years), the aforementioned norms were still applied to allow for a comparison with previously published studies. The BRIEF-SR consists of eight subscales (inhibit, shift, monitor, emotional control, working memory, plan/organize, organization of materials, and task completion), which form two index scores ("Behavioral Regulation Index" (BRI) and "Metacognition Index" (MI)) which, in turn, form one overall score ("Global Executive Composite" (GEC)). The subscale "shift" is commonly used as an indicator of the patient's flexibility in daily life. The BRIEF-SR offers scaled scores as well as *t*-scores. The latter were entered into analyses, with higher scores indicating a higher level of dysfunction (Drechsler and Steinhausen, 2013). *T*-scores at or above 65 are defined as clinically significant. The BRIEF-SR shows good internal consistency (α between .73 and .96) and good test-retest reliability (r between .78 and .86) (Drechsler and Steinhausen, 2013).

Parental self-report of executive functioning

The German version of the "Behavior Rating Inventory for Executive Functioning – Parent Form" (BRIEF-PF) (Gioia et al., 2000) consists of 86 items and aims to provide information about a child's everyday behavior related to specific domains of executive functioning. It is composed of seven subscales similar to the BRIEF-SR ("inhibit," "shift," "monitor," "emotional control," "working memory," "plan/organize," and "organization of materials") with one newly added subscale "initiate" instead of "task completion." The global scores are similar to the "BRIEF-SR" (GEC, BRI, MI). The BRIEF-PF also provides scaled scores and *t*-scores. Again, *t*-scores served as the dependent variable. Internal consistency lies between α = .75 and .89, with α = .95 for the overall score GEC. The reliability as assessed by interrater method lies between r = .35 and 60 and r = .56 for the GEC.

Statistical analyses

All analyses were conducted with SPSS 22 applying a statistical significance level of $\alpha = .05$. Four extreme outliers (more than 3 standard deviations (*SD*) from the mean) were found on the perseverative error of the WCST (boxplot) and were removed from future analyses. No other extreme outliers were found. The assumption of normality (Shapiro-Wilk test, $p < .05$) was satisfied for most scores, except for the CCI and four subscales of the BRIEF-PF (inhibit, shift, monitor, and working memory). Demographic and clinical characteristics were described via means (*M*), *SD*, minimum and maximum. There was no missing data.

To assess the relationship between the different neuropsychological measures (TMT-4, WCST, RCFT, and GEFT), the neuropsychological tests and ratings on both versions of the BRIEF as well as the self- and parental-reports, a series of Pearson correlation coefficients were calculated. Power calculations with the software package G*Power 3.1.9.2 recommend a total sample size of 38 to detect a significant large correlation ($r = .05$) with a power of 0.95. To give consideration to the violation of normality, additional nonparametric correlation coefficients (Kendall´s tau-b, τ_b) were determined. As the differences proved insubstantial and to allow for a better comparison with previous studies, only Pearson's *r* is reported. The coefficient ranges between −1 and +1 and is defined as small (.10–.29), medium (.30–.49), or large (.50–1.0). The composition of the BRIEF-SR and the BRIEF-PF is mostly comparable; however, there are three non-overlapping subscales (task completion, monitor, and initiate) that were omitted in the comparison. Nevertheless, they are still included in the summary scores (BRI, MI, and GEC). Additionally, paired *t*-tests were run to compare the magnitude of differences between the patient and parental scores on the overlapping BRIEF subscales. Six patients (15% of total patient sample) were older than the maximum age of the German norms, that is, >16.11 years, but the maximum age norms were applied. To assess the result of this, the sample was reanalyzed without these six patients, which did not alter the main findings. Effect sizes (Cohen's *d*) were also calculated and reported.

In a final step, a standard multiple regression was conducted to predict self-reported flexibility (BRIEF-SR subscale "shift" as dependent variable) from the neuropsychological tests assessing set-shifting (TMT-4 and WCST as independent variables). To detect a significant result with a large effect size following two predictors, a sample size of 48 is recommended. All statistical assumptions for multiple regression, that is, homoscedasticity of residuals, normal distribution of residuals (P-P plot), multicollinearity (tolerance >0.1, VIF <10), linearity, independence of observations (Durbin-Watson = 1.591), outliers (<3 *SD*s), high leverage points (leverage value <.02), and high influential points (Cook's distance <1), were met. Due to the explanatory nature of the study, no corrections for multiple testing were made.

Results

Demographic and clinical characteristics

The clinical characteristics of the final sample ($n = 40$) are outlined in Table 1. In total, 30 patients (75%) had a restrictive AN and 10 patients (25%) a binge/purge AN subtype. Twenty-six patients (65%) received inpatient treatment for the first time,

Table 1. Sample characteristics (N = 40).

	$M \pm SD$	Minimum	Maximum
Age (in years)	15.1 ± 1.6	11.3	17.8
BMI (kg/m^2)	15.2 ± 1.1	12.4	17.1
BMI-percentile	2.2 ± 3.3	0	10
Duration of illness (in months)	15.0 ± 12.4	3	68
Intelligence (CFT-20-R)	104.5 ± 14.9	85	136
Eating psychopathology			
EDI-2 (global score)	304.9 ± 45.1	220	388
TMT-4 (scaled score)	11.2 ± 1.5	8	14
WCST (perseveration error)	5.5 ± 7.7	0	33.3
RCFT – OCI	2.0 ± 0.6	0.7	2.7
RCFT – SI	0.9 ± 0.5	0.2	1.8
RCFT – CCI	1.1 ± 0.3	0.4	1.7
GEFT (raw score)	11.6 ± 4.4	1	18
BRIEF-SR[a]			
Inhibit	47.9 ± 7.5	35	68
Shift	58.1 ± 10.4	38	84
Emotional Control	54.6 ± 10.9	35	81
Working Memory	48.2 ± 9.7	32	71
Plan/Organize	48.9 ± 9.3	36	72
Organization of Materials	47.3 ± 9.2	37	70
BRIEF-PF			
Inhibit	45.8 ± 6.7	40	67
Shift	59.1 ± 12.6	37	92
Emotional Control	56.8 ± 10.39	40	77
Working Memory	42.7 ± 6.6	38	61
Plan/Organize	43.2 ± 7.4	36	61
Organization of Materials	44.4 ± 8.6	37	62

Note. BMI = body mass index; CFT-20-R = Culture Fair Intelligence Test-20 revised; EDI-2 = Eating Disorder Inventory 2 raw global score; TMT-4 = Trail Making Test Condition 4 in scaled scores; WCST = Wisconsin Card Sorting Test; RCFT = Rey Complex Figures Test; OCI = Order of Construction Index, range 0–3.33; SI = Style Index, range 0–2; CCI = Central Coherence Index, range 0–2; GEFT = Group Embedded Figures Test, range 0–18. BRIEF-SR, -PF = Behavior Rating Inventory of Executive Functioning – Self-Report, – Parent Form. All BRIEF scores are age and gender corrected t-scores.
[a]Norms of the German version of the BRIEF exist for children and adolescents between 11 and 16.11 years, for patients older than 16.11 years (n = 6), we applied the aforementioned norms as well.

while 11 patients (27.5%) had their second and three patients (7.5%) their third hospital stay. Eighteen patients (45%) showed a weight below the first BMI-percentile and three of these patients (7.5%) received tube feeding during their assessment. Eleven patients (27.5%) received medication, which was either antidepressants (6 patients, 15%) or neuroleptics (5 patients, 12.5%). Twenty-three patients (57.5%) had a comorbid psychiatric disorder. Of these, the patients mainly suffered from comorbid depression (18 patients, 78.3%). Three patients (13.1%) had an obsessive-compulsive disorder, one patient (4.3%) anxiety and one patient (4.3%) posttraumatic stress disorder. The overall socioeconomic status was 50.2 (SD ± 13.3), indicating a middle status.

Correlational analyses between neuropsychological measures

Pearson's r revealed small to large correlations between the neuropsychological tests (see Table 2). Significant medium and large positive correlations were found for the different indices of the RCFT (SI and OCI, SI and CCI, CCI and OCI) as well as the GEFT and CCI assessing central coherence. A medium negative correlation was found for the GEFT and WSCT.

Table 2. Pearson correlations between neuropsychological measures.

	TMT-4 scaled score	WCST perseveration error	RCFT OCI	RCFT SI	RCFT CCI	GEFT raw score
TMT-4 scaled score						
WCST perseveration error[a]						
RCFT OCI						
RCFT SI	.21		**.42****			
RCFT CCI	.22		**.79****	**.86****		
GEFT raw score	.29	**−.38***	.18	.27	**.31***	

Note. Pearson´s r is defined as small (.10 to .29), medium (.30 to .49) or large (.50 to 1.0). Correlations below .10 were omitted from the table to facilitate readability. Significant correlations are marked in bold. TMT-4 = Trail Making Test Condition 4 in scaled scores; WCST = Wisconsin Card Sorting Test; RCFT = Rey Complex Figures Test; OCI = Order of Construction Index, range 0–3.33; SI = Style Index, range 0–2; CCI = Central Coherence Index, range 0–2; GEFT = Group Embedded Figures Test, range 0–18. [a]n = 36.

Correlational analyses between neuropsychological measures and patient self-report

Results of the correlational analyses between test performance and patient self-report of executive functioning are shown in Table 3. Pearson's *r* revealed small to medium correlations. Significant medium positive correlations for central coherence were found for both SI and CCI of the RCFT and the BRIEF-SR global scores MI and GEC as well as the subscales working memory, plan/organize, and task completion. The GEFT was significantly positively correlated with the BRIEF-SR subscale "organization of materials." For set-shifting, the WCST revealed significant medium negative correlations between the BRIEF-SR global score GEC and the subscales "shift," "emotional control," and "monitor."

Correlational analyses between neuropsychological measures and parental self-report

The parental self-report and the adolescent's test performance revealed mostly small and few medium associations. Significant correlations were found for set-shifting only, with the WCST showing medium negative correlations between the BRIEF-PF global scores BRI, MI and GEC as well as the subscales shift, emotional control, and monitor (see Table 4).

Correlational analyses between patient and parental self-report

The parental version of the BRIEF was mainly filled out by the patients' mothers (30 patients, 75%), for nine patients (22.5%) the father and for one patient (2.5%) a close family member answered the questionnaire. The correlational analyses between the two self-reports was conducted for the overlapping subscales ("inhibit," "shift," "emotional control," "working memory," "plan/organize," and "organization of materials") only. Pearson correlations revealed small to medium and one large correlation for the BRIEF-

Table 3. Pearson correlations between neuropsychological measures and patient self-report (N = 40).

	BRIEF-SR										
	BRI	MI	GEC	Inhibit	Shift	Emotional Control	Monitor	Working Memory	Plan/Organize	Organization of Materials	Task Completion
TMT-4 scaled score				.11							−.20
WCST perseveration error[a]	**−.42****		−.22	−.24	**−.33***	**−.38***	**−.39***	.13		−.18	
RCFT OCI	.17	.24	.23	.26	.14		.27	.15	.21	.20	.22
RCFT SI	.16	**.46****	**.35***	.14	.11	.17	.23	**.37***	**.38***	.26	**.39***
RCFT CCI	.17	**.43****	**.34***	.24	.13		.26	**.32***	**.36***	.29	**.35***
GEFT raw score		.11		.12	.10					**.31***	

Note. All BRIEF-SR scores are age and gender corrected t-scores. Pearson´s r is defined as small (.10 to .29), medium (.30 to .49) or large (.50 to 1.0). Correlations below .10 were omitted from the table to facilitate readability. Significant correlations are marked in bold. TMT-4 = Trail Making Test Condition 4 in scaled scores; WCST = Wisconsin Card Sorting Test; RCFT = Rey Complex Figures Test; OCI = Order of Construction Index, range 0–3.33; SI = Style Index, range 0–2; CCI = Central Coherence Index, range 0–2; GEFT = Group Embedded Figures Test, range 0–18; BRIEF-SR = Behavior Rating Inventory of Executive Functioning – Self-Report; BRI = Behavioral Regulation Index; MI = Metacognition Index; GEC = Global Executive Composite. [a]n = 36. *Correlation is significant at the .05 level (two-tailed). **Correlation is significant at the .01 level (two-tailed).

Table 4. Pearson correlations between neuropsychological measures and parental self-report (N = 40).

	BRIEF-PF										
	BRI	MI	GEC	Inhibit	Shift	Emotional Control	Initiate	Working Memory	Plan/ Organize	Organization of Materials	Monitor
TMT-4 scaled score			.11	.23		.10	−.14	.10	.10		.17
WCST perseveration error[a]	**−.46****	**−.39***	**−.42***	.20	**−.40***	**−.38***	−.28		.25		**−.45****
RCFT OCI		−.22		−.29	.14			−.15		.15	
RCFT SI							−.12	−.11	.27	.25	
RCFT CCI		−.11		−.22				−.18	.12	.25	
GEFT raw score		.11								.20	

Note. All BRIEF-SR scores are age and gender corrected t-scores. Pearson´s r is defined as small (.10 to .29), medium (.30 to .49) or large (.50 to 1.0). Correlations below .10 were omitted from the table to facilitate readability. Significant correlations are marked in bold. TMT-4 = Trail Making Test Condition 4 in scaled scores; WCST = Wisconsin Card Sorting Test; RCFT = Rey Complex Figures Test; OCI = Order of Construction Index, range 0–3.33; SI = Style Index, range 0–2; CCI = Central Coherence Index, range 0–2; GEFT = Group Embedded Figures Test, range 0–18; BRIEF-PF = Behavior Rating Inventory of Executive Functioning – Parent Form; BRI = Behavioral Regulation Index; MI = Metacognition Index; GEC = Global Executive Composite. [a]n = 36. **Correlation is significant at the .05 level (two-tailed). **Correlation is significant at the .01 level (two-tailed).*

SR subscale "working memory" and the BRIEF-PF subscale "plan/organize." Significant medium positive correlations were found for the BRIEF-SR subscale "shift" and the BRIEF-PF subscale "plan/organize"; the same was found for the BRIEF-SR subscale "working memory" and the BRIEF-PF subscales "plan/organize" and "organization of materials" as well as the BRIEF-SR subscale "organization of materials" and the corresponding BRIEF-PF subscale (see Table 5).

Overall, the BRIEF scores of both the patients and their parents fell within the normal range (t-score < 65) (Drechsler and Steinhausen, 2013). For details see Table 1. Additionally, conducted paired t-tests revealed a statistically significant mean difference between the patient and parental rating on the subscales "working memory" ($t(39) = 3.113$, $p = .003$, $d = .663$) and "plan/organize" ($t(39) = 3.570$, $p = .001$, $d = .705$), with the adolescents scoring significantly higher than their parents.

Standard multiple regression analyses

A standard multiple regression was run to examine whether day-to-day flexibility (operationalized via the BRIEF-SR subscale "shift") can be predicted from neuropsychological performance on the two measures examining flexibility. The TMT-4 (scaled score) and WCST (perseveration error) were selected as predictors. The regression model was nonsignificant ($F(2,35) = 2.052$, $p = .145$, $R^2 = .11$), indicating that the test performance accounted for only 11% of the explained variability. In summary, neuropsychological test performance did not significantly predict self-rated flexibility in the context of everyday-life situations.

Discussion

The present study aimed to address the relationship between different neuropsychological measures, as well as the relationship between neuropsychological tests and self- and parental-report findings of executive functioning in adolescents with acute AN, a field where more research is needed. In addition, associations between patient and

Table 5. Pearson correlations between patient and parental self-report ($N = 40$).

		BRIEF-PF					
		Inhibit	Shift	Emotional Control	Working Memory	Plan/ Organize	Organization of Materials
BRIEF-SR	Inhibit		.14				
	Shift	.16	.31	.29		**.35***	
	Emotional Control	.27		.25		.28	
	Working Memory	.22		.14		**.53****	**.39***
	Plan/Organize	.13				.29	.12
	Organization of Materials		−.27	−.15		.13	**.45****

*Note. All correlations are based on age and gender corrected t-scores. Pearson´s r is defined as small (.10 to .29), medium (.30 to .49) or large (.50 to 1.0). Correlations below .10 were omitted from the table to facilitate readability. Significant correlations are marked in bold. BRIEF-SR, -PF = Behavior Rating Inventory of Executive Functioning – Self-Report, – Parent Form. The subscales initiate, monitor and task completion as well as the summary scores were omitted, as they do not overlap in both BRIEF-SR and BRIEF-PF. *Correlation is significant at the .05 level (two-tailed). **Correlation is significant at the .01 level (two-tailed).*

parental self-report ratings were examined. This study substantiates and extends the findings from Stedal and Dahlgren (2015), which, to date, is the only study exploring the relationship between cognitive and behavioral aspects of executive functioning in adolescent AN.

Overall, our results indicate small to large correlations between the different neuropsychological tests and low to moderate correlations between neuropsychological performance and everyday-life functioning as rated by the AN patients and their parents. Concerning the association between patient and parental self-report, our results also reveal modest agreement. The patients scored significantly higher than their parents on two subscales, though both ratings fell within the range of the normative mean and do not speak for clinically significant dysfunction in self- and parental-reported executive functioning in daily life. Moreover, despite being slightly underpowered, test performance, that is, assessing flexibility, seems to not predict the self-rated experience of flexibility in everyday life. In terms of the debate of whether or not adolescents with AN show inefficiencies in neuropsychological functioning to the same extent that adult patients do, the integration of our findings is limited, mainly due to the lack of age and gender corrected norms.

With reference to the mostly small and few large correlations between the different neuropsychological measures, a possible explanation could be that although two measures each were used for set-shifting (WSCT, TMT-4) and central coherence (RCFT, GEFT), these instruments may assess different aspects of the same construct and may differ in content and complexity (Wu et al., 2014). The WCST, for instance, is often described as a rather complicated task that involves various cognitive operations, such as inhibition, problem solving and working memory, besides the target component of set-shifting (Miyake et al., 2000, Tchanturia et al., 2012). Scoring procedures also differ with number of errors for the WCST and reaction time for the TMT-4, which might provide part of the explanation for the low correlations. For central coherence, the correlations were slightly higher, although the aforementioned limitations also apply. The medium negative relationship between WCST and GEFT is in line with the expectations; the ability to flexibly shift attention between different tasks and changing demands (high levels of set-shifting) is associated with higher scores on the GEFT, indicative of a more field-independent approach. That is, patients showing a bias toward detail are better able to restructure tasks and remove distractions as well as rapidly differentiate between visual stimuli. However, due to the lack of comparable studies and the small sample size, we are limited in making further interpretations.

In terms of the associations between test results and patient self-report of executive functioning, our findings reveal significant medium negative associations between measures of set-shifting (WCST) and several subscales of the BRIEF-SR (see Table 3). This suggests that more problems in perseverative responding are associated with less self-rated problems in shifting flexibly from one situation to another (subscale “shift”), regulating emotional responses (subscale “emotional control”) and being aware of one´s own effect on others (subscale “monitor”), which seems contrary to the assumptions. Measures of central coherence (SI and CCI), in contrast, are positively correlated with other BRIEF-SR subscales. This emphasizes that more global processing strategies and a tendency to continuity are associated with increased skills in holding information when completing a task (subscale “working memory”), anticipating future events (subscale

"plan/organize") and finishing tasks in a timely fashion (subscale "task completion"), which is a relationship in the expected direction. For the parental ratings and the test scores, the correlations are lower than those from the patient self-report. Only the WCST reveals medium negative associations between the global indices and three subscales that are in line with the patient ratings ("shift," "emotional control," and "monitor"). The abovementioned low to moderate associations are in accordance with previous research in children and adolescents with traumatic brain injury (Vriezen and Pigott, 2002, Wilson et al., 2011), brain disease such as focal frontal lesions or early treated hydrocephalus(Anderson et al., 2002), ADHD (Barkley, 1991), attention and learning problems (McAuley et al., 2010), and psychosis (Niendam et al., 2007). With regard to AN, Stedal and Dahlgren (2015) also found little agreement between neuropsychological measures and self-report of executive functioning. Again, as there is only little research in patients with AN available, the generalizability of our findings is limited.

The lack of associations between neuropsychological tests and behavioral ratings could partly be explained by differences in the conception and operationalization of the measures. To begin with, these measures assess different aspects of the same underlying construct, that is, behavioral and cognitive components of executive functioning (Anderson et al., 2002, Jurado and Rosselli, 2007, Toplak, West, & Stanovich, 2013, Wilson et al., 2011). While neuropsychological assessment focusses on the underlying skills, behavioral rating measures assess the application of those skills in real-life situations, which lead to different results accordingly (McAuley et al., 2010). Concerning the operationalization, performance-based tests and rating measures also differ in terms of what outcome is reinforced, that is, the intention to capture an optimal performance at a specific point of time (with the help of instructions that maximize the output and the definition of distinct goals) versus the assessment of typical behavior and difficulties in everyday-life over a longer period of time (Toplak et al., 2013). Other reasons refer to the development and administration of the measures, as most neuropsychological tests have not been empirically validated to predict real-life behavior but to measure the impairment of functions and capacities of the frontal lobe (Jurado and Rosselli, 2007). As a result, the highly controlled laboratory setting may not be sensitive enough to capture day-to-day failures. As such, AN patients may have difficulties in their daily life but function well on assessment measures in a standardized test environment (Chevignard, Soo, Galvin, Catroppa, & Eren, 2012, Stedal and Dahlgren, 2015). In line with this, the test setting has only few similarities with the real world, as distractions are minimized and the patient is given stepwise instructions (Sbordone, 1996). Following this, the laboratory setting may be unsuitable to detect everyday impairment that becomes more apparent in complex situations (Burgess, Alderman, Evans, Emslie, & Wilson, 1998, Gioia and Isquith, 2004). In addition, the neuropsychological test needs to be precisely tailored to a child's (dys-) functioning, as tasks that are either too simple or too complex may not be able to identify difficulties. This, however, seems questionable, as the majority of measures were validated in adults and may be more sensible to deficits in this age group than in adolescents (Vriezen and Pigott, 2002). Performance-based tests also rely on a wide range of skills that aggravate the isolation of specific processes (Jurado and Rosselli, 2007). Overall, neuropsychological tests and behavioral ratings offer different aspects of

the same construct and cannot be used as a proxy for each other. With regard to the assessment of cognitive functioning in adolescents with AN, our findings suggest the need for a combination of standardized tests with self-reports of executive functioning as well as the use of multiple informants to ensure a comprehensive view.

The low associations between self- and parental-ratings of executive functioning also allow for several interpretations. Firstly, the patients and their parents share different perspectives concerning a certain behavior, as patients have more insight about how their behavior is motivated and influenced when compared to an observation from the outside (Anderson et al., 2002). Secondly and referring to attribution theory, parental ratings may focus more closely on the person, whereas the patient more often refers to the situation (internal versus external attribution; Parkinson, 2007). Thirdly, differences in the degree of insight as well as social desirability might have an influence on the ratings. Fourthly, family stress and parental dysfunction may influence the parent-adolescent correspondence, for instance due to increased irritation and reduced tolerance for the adolescent behavior (Dimler, Natsuaki, Hastings, Zahn-Waxler, & Klimes-Dougan, 2016, Kolko and Kazdin, 1993). Differences could also occur between the judgments of the mothers and fathers and their children; available literature points toward more mother-adolescent-agreement (Seiffge-Krenke and Kollmar, 1998). However, as these variables have not been thoroughly investigated in the current study, our results should be interpreted with some caution.

This study shows strengths and weaknesses. To begin with the strengths, the study examined the association between neuropsychological assessment and rating scales in adolescent AN, where, to date, little research is available. Beyond that, we evaluated adolescents between 11 and 17.11 years, representing the age group with the highest AN incidence rates (Herpertz-Dahlmann, 2015). Our sample also included a well-defined patient group after an extensive screening process. Moreover, this study includes data from different measures and sources that are widely used in AN research and add to a more comprehensive assessment of executive functioning. The BRIEF also counts as one of the most commonly used rating scales of executive functioning in adolescents. Weaknesses mainly refer to the lack of a healthy control group, which restricts the generalizability of our results. The relatively small sample size forms another limitation, as the results could partly be attributable to a lack of statistical power. The representativeness of our data is also slightly limited, as German norms of the BRIEF-SR are only available up to 16.11 years, which did not apply for six patients (15%) who were marginally older. However, a reanalysis of the data without these six patients did not lead to different findings. Another limitation for the representativeness of our data is caused by the fact that no German norms are available for the D-KEFS TMT. Collecting psychiatric comorbidities solely from the patients' medical records without a more structured assessment can be regarded as another limitation. Moreover, the presence of psychiatric comorbidities in our sample could suggest that the current results are not specific to AN. However, pure AN is rarely encountered in clinical practice and the rate of psychiatric comorbidities is high with 47–60% of patients, for example, depression is diagnosed in 35–39% of cases (Bühren et al., 2014, Jaite et al., 2013).

Another influencing factor that could affect the results is the age range of the sample, as younger patients might have more problems self-rating their behavior than older subjects. Similarly, the parental ratings may also differ as parents of younger patients

may have a closer relationship when compared to older patients who have developed more independence. The same applies for the duration of illness and the number and severity of comorbid symptoms, which could result in more self-rated problems in everyday functioning for both the patients and their parents when bearing in mind that comorbidity is associated with a higher level of distress and problems in psychosocial adaptation (Herpertz-Dahlmann et al., 2001). However, these aspects also apply for comparable studies in clinical samples.

Taken together, the current findings suggest that performance-based and self-rating measures provide different types of information that are each associated with advantages and disadvantages. The moderate relationship between the measures demonstrates that the currently available tests show some kind of ecological validity; however, the predictive power of a standardized assessment with regard to cognitive functioning as it appears in the daily-life of adolescents with AN needs to be questioned and the generalizability of data obtained in a laboratory is limited (Chaytor and Schmitter-Edgecombe, 2006). Especially for younger patients with AN, where there are conflicting results in terms of neuropsychological functioning, to solely rely on performance-based measures is debatable. Nonsignificant test results also do not rule out possible difficulties in everyday-life functioning and vice versa. With respect to clinical relevance, the current findings therefore emphasize the importance of self-ratings measures, preferably from patients and important others. Consequently, it can be concluded that so far the best approach to accurately assess cognitive inefficiencies and difficulties in everyday-life functioning in adolescents with AN is a comprehensive assessment that combines data from multiple measures and sources with the results of standardized tests (Chevignard et al., 2012).

Disclosure statement

The authors report no conflict of interest.

Funding

No specific funding was utilized in conducting this study or preparing this manuscript, which was produced in partial fulfillment of PhD requirements for the first author.

ORCID

Betteke Maria van Noort http://orcid.org/0000-0002-9768-3383

References

Agras, W. S., Crow, S., Mitchell, J. E., Halmi, K. A., and Bryson, S. (2009). A 4-year prospective study of eating disorder NOS compared with full eating disorder syndromes. *The International Journal of Eating Disorders*, *42*(6), 565–570.

Allen, K. L., Byrne, S. M., Hii, H., van Eekelen, A., Mattes, E., and Foster, J. K. (2012). Neurocognitive functioning in adolescents with eating disorders: A population-based study. *Cognitive Neuropsychiatry*, *18*(5), 355–375.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, Forth Edition, Text Revision (DSM-IV-TR)*. Washington, DC: Author.

Anderson, V., Anderson, P., Northam, E., Jacobs, R., and Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology*, *8*(4), 231–240.

Andres-Perpina, S., Lozano-Serra, E., Puig, O., Lera-Miguel, S., Lazaro, L., and Castro-Fornieles, J. (2011). Clinical and biological correlates of adolescent anorexia nervosa with impaired cognitive profile. *European Child & Adolescent Psychiatry*, *20*(11–12), 541–549.

Arcelus, J., Haslam, M., Farrow, C., and Meyer, C. (2013). The role of interpersonal functioning in the maintenance of eating psychopathology: A systematic review and testable model. *Clinical Psychology Review*, *33*(1), 156–167.

Arcelus, J., Mitchell, A. J., Wales, J., and Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders. *Archives of General Psychiatry*, *68*(7), 724–731.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *The Journal of Abnormal Child Psychology*, *19*(2), 149–178.

Bentz, M., Jepsen, J. R., Kjaersdam Telleus, G., Moslet, U., Pedersen, T., Bulik, C. M., and Plessen, K. J. (2017). Neurocognitive functions and social functioning in young females with recent-onset anorexia nervosa and recovered individuals. *Journal Eating Disorders*, *5*, 5.

Blakemore, S. J., Burnett, S., and Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping*, *31*(6), 926–933.

Booth, R. (2006). *Local-global processing and cognitive style in autism spectrum disorders and typical development*. London: Institute of Psychiatry, Kings's College London.

Bühren, K., Schwarte, R., Fluck, F., Timmesfeld, N., Krei, M., Egberts, K., … Herpertz-Dahlmann, B. (2014). Comorbid psychiatric disorders in female adolescents with first-onset anorexia nervosa. *European Eating Disorders Review*, *22*(1), 39–44.

Burgess, P. W., Alderman, N., Evans, J. J., Emslie, H., and Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society : JINS*, *4*, 547–558.

Chaytor, N., and Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*(4), 181–197.

Chaytor, N., and Schmitter-Edgecombe, M. (2006). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*(4), 181–196.

Chevignard, M. P., Soo, C., Galvin, J., Catroppa, C., and Eren, S. (2012). Ecological assessment of cognitive functions in children with acquired brain injury: A systematic review. *Brain Injury*, *26*(9), 1033–1057.

Currin, L., Schmidt, U., Treasure, J., and Jick, H. (2005). Time trends in eating disorder incidence. *The British Journal of Psychiatry: The Journal of Mental Science*, *186*, 132–135.

Dahlgren, C. L., Lask, B., Landro, N. I., and Rø, Ø. (2014). Patient and parental self-reports of executive functioning in a sample of young female adolescents with anorexia nervosa before and after cognitive remediation therapy. *European Eating Disorders Review*, *22*(1), 45–52.

Danner, U. N., Sanders, N., Smeets, P. A., van Meer, F., Adan, R. A., Hoek, H. W., and 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.

Davies, H., and Tchanturia, K. (2005). Cognitive remediation therapy as an intervention for acute anorexia nervosa: A case report. *European Eating Disorders Review*, *13*(5), 311–316.

Delis, D., Kaplan, E., and Kramer, J. (2001). *Delis–Kaplan executive function system (D–KEFS)*. San Antonio, Texas: Psychological Corporation.

Dimler, L. M., Natsuaki, M. N., Hastings, P. D., Zahn-Waxler, C., and Klimes-Dougan, B. (2016). Parenting effects are in the eye of the beholder: Parent-adolescent differences in perceptions

affects adolescent problem behaviors. *Journal of Youth and Adolescence*. doi:10.1007/s10964-016-0612-1

Drechsler, R., and Steinhausen, H.-C. (2013). *BRIEF. Verhaltensinventar zur Beurteilung exekutiver Funktionen*. Göttingen: Hogrefe.

Drühe-Wienholt, C.-M., and Wienholt, W. (2004). *Computergestütztes Kartensortierverfahren. Manual*. Frankfurt/Main: Harcourt Test Services.

Fichter, M., and Quadflieg, N. (1999). *Strukturiertes Inventar für Anorektische und Bulimische Essstörungen (SIAB-EX). Handanweisung*. Göttingen: Hogrefe-Verlag für Psychologie.

Fichter, M., and Quadflieg, N. (2001). Das Strukturierte Interview für Anorektische und Bulimische Ess-Störungen nach DSM-IV und ICD-10 zur Expertenbeurteilung (SIAB-EX) und dazugehöriger Fragebogen zur Selbsteinschätzung (SIAB-S). *Verhaltenstherapie*, *11*, 314–325.

Fitzpatrick, K. K., Darcy, A., Colborn, D., Gudorf, C., and Lock, J. (2012). Set-shifting among adolescents with anorexia nervosa. *The International Journal of Eating Disorders*, *45*(7), 909–912.

Gioia, G. A., and Isquith, P. K. (2004). Ecological assessment of executive function in traumatic brain injury. *Developmental Neuropsychology*, *25*(1–2), 135–158.

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology*, *6*(3), 235–238.

Giombini, L., Moynihan, J., Turco, M., and Nesbitt, S. (2016). Evaluation of individual cognitive remediation therapy (CRT) for the treatment of young people with anorexia nervosa. *Eating and Weight Disorders*. doi:10.1007/s40519-016-0322-4

Godart, N. T., Perdereau, F., Rein, Z., Berthoz, S., Wallier, J., Jeammet, P., and Flament, M. F. (2007). Comorbidity studies of eating disorders and mood disorders. Critical review of the literature. *Journal of Affective Disorders*, *97*(1–3), 37–49.

Guy, S. C., Isquith, P. K., and Gioia, G. A. (2004). *BRIEF-SR: Behavior rating inventory of executive function – Self-report version. professional manual*. Lutz, FL: Psychological Assessment Resources.

Heaton, R., Chelune, G., Kay, G., and Curtiss, G. (1993). *Wisconsin card sorting test manual: Revised and explanded*. Odessa, Florida: Psychological Assessment Resources, Inc.

Herpertz-Dahlmann, B. (2015). Adolescent eating disorders: Update on definitions, symptomatology, epidemiology, and comorbidity. *Child and Adolescent Psychiatric Clinics of North America*, *24*(1), 177–196.

Herpertz-Dahlmann, B., Müller, B., Herpertz, S., Heussen, N., Hebebrand, J., and Remschmidt, H. (2001). Prospective 10-year follow-up in adolescent anorexia nervosa - course, outcome, psychiatric comorbidity, and psychosocial adaptation. *Journal of Child Psychology and Psychiatry*, *42*(5), 603–612.

Hoek, H. W., and van Hoeken, D. (2003). Review of the prevalence and incidence of eating disorders. *The International Journal of Eating Disorders*, *34*(4), 383–396.

Hollingshead, A. (1975). Four Factor Index of Social Status. *unpublished working paper*.

Holtkamp, K., and Herpertz-Dahlmann, B. (2005). Anorexia und Bulimia nervosa im Kindes- und Jugendalter. *Deutsches Arzteblatt*, *102*(1–2), 50–58.

Jaite, C., Hoffmann, F., Glaeske, G., and Bachmann, C. J. (2013). Prevalence, comorbidities and outpatient treatment of anorexia and bulimia nervosa in German children and adolescents. *Eating and Weight Disorders*, *18*(2), 157–165.

Jurado, M. B., and Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233.

Kanakam, N., Raoult, C., Collier, D., and Treasure, J. (2013). Set shifting and central coherence as neurocognitive endophenotypes in eating disorders: A preliminary investigation in twins. *The World Journal of Biological Psychiatry*, *14*(6), 464–475.

Katzman, D. K. (2005). Medical complications in adolescents with anorexia nervosa: A review of the literature. *The International Journal of Eating Disorders*, *37*(Suppl), S52–59. discussion S87-59.

Kaye, W. H., Fudge, J. L., and Paulus, M. (2009). New insights into symptoms and neurocircuit function of anorexia nervosa. *Nature Reviews. Neuroscience*, *10*(8), 573–584.

Kidd, A., and Steinglass, J. (2012). What can cognitive neuroscience teach us about anorexia nervosa? *Current Psychiatry Reports*, *14*(4), 415–420.

Kjaersdam Telleus, G., Jepsen, J. R., Bentz, M., Christiansen, E., Jensen, S. O., Fagerlund, B., and Thomsen, P. H. (2015). Cognitive profile of children and adolescents with anorexia nervosa. *European Eating Disorders Review*, *23*(1), 34–42.

Kolko, D., and Kazdin, A. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among children, parent and teacher reports. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *34*(6), 991–1006.

Kromeyer-Hauschild, K., Wabitsch, M., Kunze, D., Geller, F., Geiß, H., Hesse, V., … Hebebrand, J. (2001). Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschrift Kinderheilkunde*, *149*, 807–818.

Lang, K., Lloyd, S., Khondoker, M., Simic, M., Treasure, J., and Tchanturia, K. (2015). Do children and adolescents with anorexia nervosa display an inefficient cognitive processing style? *PLoS ONE*, *10*(7), e0131724.

Lang, K., Lopez, C., Stahl, D., Tchanturia, K., and 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.

Lang, K., Stahl, D., Espie, J., Treasure, J., and Tchanturia, K. (2014). Set shifting in children and adolescents with anorexia nervosa: An exploratory systematic review and meta-analysis. *The International Journal of Eating Disorders*, *47*(4), 394–399.

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

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., … Denckla, M. B. (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*, 643–662.

Mahone, E. M., Zabel, T. A., Levey, E., Verda, M., and Kinsman, S. (2002). Parent and self-report ratings of executive function in adolescents with myelomeningocele and hydrocephalus. *Child Neuropsychology*, *8*(4), 258–270.

McAnarney, E. R., Zarcone, J., Singh, P., Michels, J., Welsh, S., Litteer, T., … Klein, J. D. (2011). Restrictive anorexia nervosa and set-shifting in adolescents: A biobehavioral interface. *The Journal of Adolescent Health*, *49*(1), 99–101.

McAuley, T., Chen, S., Goos, L., Schachar, R., and Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society : JINS*, *16*(3), 495–505.

McCandless, S., and O'Laughlin, L. (2007). The clinical utility of the behavior rating inventory of executive function (BRIEF) in the diagnosis of ADHD. *Journal of Attention Disorders*, *10*(4), 381–389.

Meyers, J., and Meyers, K. (1995). *Rey complex figure and recognition trial: Professional manual*. Odessa, Florida: Psychological Assessment Resources.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Niendam, T. A., Horwitz, J., Bearden, C. E., and Cannon, T. D. (2007). Ecological assessment of executive dysfunction in the psychosis prodrome: A pilot study. *Schizophrenia Research*, *93*(1–3), 350–354.

Parkinson, B. (2007). Soziale Wahrnehmung und Attribution. *In* K. Jonas, W. Stroebe, & M. Hewstone (Eds.), *Sozialpsychologie* (pp. 70–112). Heidelberg: Springer.

Paul, T., and Thiel, A. (2004). *Eating disorder inventory-2*. Göttingen: Hogrefe.

Remschmidt, H., Schmidt, M. H., and Poustka, F. (2006). *Multiaxiales Klassifikationsschema für psychische Störungen des Kindes- und Jugendalters nach ICD-10 der WHO. Mit einem synoptischen Vergleich von ICD-10 und DSM-IV (5. vollständige überarbeitete Auflage)*. Bern: Huber.

Sbordone, R. J. (1996). *Ecological validity: Some critical issues for the neurologist*. England: Gr Press/St. Lucie Press.

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, *4*(2), 419–436.

Seiffge-Krenke, I., and Kollmar, F. (1998). Discrepancies between mothers′ and fathers′ perceptions of sons′ and daughters′ problem behavior: A longitudinal analysis of parent-adolescent agreement of internalising and externalising problem behavior. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *39*(5), 687–697.

Shott, M. E., Filoteo, J. V., Bhatnagar, K. A., Peak, N. J., Hagman, J. O., Rockwell, R., … Frank, G. K. (2012). Cognitive set-shifting in anorexia nervosa. *European Eating Disorders Review*, *20* (5), 343–349.

Stedal, K., and Dahlgren, C. L. (2015). Neuropsychological assessment in adolescents with anorexia nervosa - exploring the relationship between self-report and performance-based testing. *Journal Eating Disorders*, *3*, 27.

Talbot, A., Hay, P., Buckett, G., and Touyz, S. (2015). Cognitive deficits as an endophenotype for anorexia nervosa: An accepted fact or a need for re-examination? *The International Journal of Eating Disorders*, *48*(1), 15–25.

Tchanturia, K., Davies, H., Roberts, M., Harrison, A., Nakazato, M., Schmidt, U., … Morris, R. (2012). Poor cognitive flexibility in eating disorders: Examining the evidence using the Wisconsin card sorting task. *PLoS ONE*, *7*(1), e28331.

Tchanturia, K., Lounes, N., and Holttum, S. (2014). Cognitive remediation in anorexia nervosa and related conditions: A systematic review. *European Eating Disorders Review*, *22*(6), 454–462.

Tenconi, E., Santonastaso, P., Degortes, D., Bosello, R., Titton, F., Mapelli, D., and Favaro, A. (2010). Set-shifting abilities, central coherence, and handedness in anorexia nervosa patients, their unaffected siblings and healthy controls: Exploring putative endophenotypes. *The World Journal of Biological Psychiatry*, *11*, 813–823.

Toplak, M. E., Bucciarelli, S. M., Jain, U., and Tannock, R. (2009). Executive functions: Performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*(1), 53–72.

Toplak, M. E., West, R. F., and Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*(2), 131–143.

van Noort, B. M., Kraus, M. K. A., Pfeiffer, E., Lehmkuhl, U., and Kappel, V. (2015). Neuropsychological and behavioural short-term effects of cognitive remediation therapy in adolescent anorexia nervosa: A pilot study. *European Eating Disorders Review*, *24*, 69–74.

van Noort, B. M., Pfeiffer, E., Ehrlich, S., Lehmkuhl, U., and Kappel, V. (2016). Cognitive performance in children with acute early-onset anorexia nervosa. *European Child & Adolescent Psychiatry*, *25*, 1233–1244.

Vriezen, E. R., and Pigott, S. E. (2002). The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology*, *8*(4), 296–303.

Weiß, R. H. (2006). *Grundintelligenztest Skala 2 - Revision (CFT 20-R)*. Göttingen: Hogrefe.

Wilson, K. R., Donders, J., and Nguyen, L. (2011). Self and parent ratings of executive functioning after adolescent traumatic brain injury. *Rehabilitation Psychology*, *56*(2), 100–106.

Witkin, H., Oltman, P., Raskin, E., and Karp, S. (1971). *A manual for the embedded figures tests*. Palo Alto, CA: Consulting Psychologists Press.

Wu, M., Brockmeyer, T., Hartmann, M., Skunde, M., Herzog, W., and Friederich, H. C. (2014). Set-shifting ability across the spectrum of eating disorders and in overweight and obesity: A systematic review and meta-analysis. *Psychological Medicine*, *44*(16), 3365–3385.