Cognitive and executive functions in anorexia nervosa
ten years after onset of eating disorder

I. Carina Gillberg,1 Maria Råstam,1 Elisabet Wentz,2 and
Christopher Gillberg1

1Child and Adolescent Psychiatry, Göteborg University, Goteborg, Sweden
2Child and Adolescent Psychiatry, Göteborg University, The Swedish National Healthcare and
Sciences, Sweden

In a longitudinal study, the authors explore the course of general cognition in anorexia nervosa (AN) over time and compare general cognitive problems, executive function deficits, attentional problems and visuomotor dysfunctions across AN individuals and healthy controls. A community-based sample of adolescent onset AN cases ($n = 40–47$) was contrasted with an age-, sex- and school matched comparison group ($n = 47–51$) on the Wechsler Adult Intelligence Scale-Revised, the Wisconsin Card Sorting Test and Luria word recall test at a mean age of 24 years. Only two of the cases tested were underweight at the time of the study. The Wechsler scale had also been administered when the groups had a mean age of 21 years. There were few differences across the two groups even though the comparison group performed significantly better on the Object Assembly subtest of the WAIS-R. IQ increased slightly but significantly over time in both groups. There was no relationship between level of starvation and poor results on tests in the AN group. A subgroup of the subjects had autism spectrum disorders. In this subgroup there were cases with test profiles similar to those observed in autism and Asperger syndrome, just as there had been on testing three years previously. Ten years after AN onset, the former AN cases showed no major neuropsychological deficits. A subgroup with autistic features had test profiles similar to those observed in autism spectrum disorders. The AN group as a whole showed poor results on the object assembly subtest indicating weak central coherence with a tendency to focus on details at the expense of configural information. This cognitive style may account for their obsession with details, with implications for psychoeducational approaches in treatment programmes/interventions.

INTRODUCTION

Some years ago we published data from a prospective six-year follow-up study of a community-based sample with teenage onset anorexia nervosa (AN) and a comparison group (Gillberg et al., 1996). Findings indicated that overall IQ — on the Wechsler Adult Intelligence Scale-Revised (WAIS-R) — was similar to that in the general population. A few other studies have shown IQ to be in the normal range (Dura & Bornstein, 1989; Kowalski, 1986), but at least two reports found high IQ in AN (Blanz, Detzner, Lay, Rose, & Schmidt, 1997; Dally, 1969). Only our study referred to a community-based group with AN and a longitudinally followed community comparison sample. The findings of the other studies could have been caused by the selective nature of the clinical samples included or the failure to include a non-AN comparison group.

There have been indications that some aspects of attentional (particularly visuospatial but also auditory) functions might be impaired in AN (Gillberg, Gillberg, Råstam, & Johansson, 1996; Green, Elliman, Wakeling, & Rogers, 1996; Hamsher, Halmi, & Benton, 1981; Horne, Van Vactor, & Emerson, 1991; Kingston,
Szmukler, Robinson, & Tress, 1996; Laessle, Krieg, & Pirke, 1989; Lauer, Gorzewski, Gerlinghoff, Backmund, & Zihl, 1999; Palazidou, Robinson, & Lishman, 1990; Seed, Dixon, McCluskey, & Young, 2000; Szmukler et al., 1992). One recent controlled study (Tchanturia et al., 2004a) found inflexibility, with problems in set-shifting tasks in AN (and slightly different set-shifting problems in BN). Some authors have attempted to link these findings to disturbed body image and parietal area dysfunction (Grunwald et al., 2001; Tomasino, 1996; Uher et al., 2005) or to hemispheric differences in AN women when judging body pictures (Smeets & Kosslyn, 2001). Our group (Gillberg et al., 1996) suggested that the failure on the Object Assembly subtest of the WAIS-R might be a consequence of over-reliance on detail and a decreased drive for central coherence (Happe, 1994). Such problems are believed to be important in autism spectrum disorders (ASD) (Jolliffe & Baron-Cohen, 1999). Interestingly, it appears that a subgroup of individuals with AN is affected by ASD (Wentz Nilsson, Gillberg, Gillberg, & Råstam, 1999). In our previous study, this subgroup showed particularly poor results on the Object Assembly subtest, and on another subtest believed to reflect autism spectrum problems, (the Picture Arrangement scale) of the WAIS-R (Gillberg et al., 1996).

Studies examining the effect of refeeding on neuropsychological outcome measures have yielded partly contradictory results. For instance, in a controlled study of 12 patients with AN followed over a 12-week period, there were deficits in memory, reaction time and motor speed which were not positively affected by weight gain (Green et al., 1996). In a controlled study of 30 AN patients, where half of the sample had recovered from AN, both the AN and the recovered subjects had impaired executive function in terms of set-shifting tasks (Tchanturia, Morris, Surguladze, & Treasure, 2002). The results were later confirmed in another AN sample before and after weight gain (Tchanturia et al., 2004b). The problems with set-shifting were linked to childhood perfectionism in that study. In yet another study, 12 patients with AN and 14 patients with bulimia nervosa had deficient attention and problem-solving (Lauer et al., 1999). After 7 months, both groups had improved cognitively in a similar fashion. In a controlled study of 46 in-patients with AN, those who had gained at least 10% of their body weight were retested on neuropsychological measures and were rescanned on Magnetic Resonance Imaging (MRI) some time after initial assessments. The AN group performed significantly worse than controls on tasks measuring attention, visuospatial ability and memory. Some, but not all, radiological measures of cerebral atrophy had normalized. There were only weak relationships between brain changes and cognitive impairments (Kingston et al., 1996).

In our six-year follow-up of a teenage onset AN group and their comparison cases we found few differences across the two groups (Gillberg et al., 1996). There was, however, a significant trough on the WAIS-R subtest of Object Assembly in the AN group. Poor psychosocial functioning was seen in a substantial subgroup of the former AN cases. One year after the six-year follow-up almost half the AN group had been examined with SPECT (Råstam et al., 2001). In spite of normalized weight, regional blood flow was significantly reduced in the AN group, and most pronounced in temporal, parietal, occipital and orbitofrontal regions.

In the light of previous findings, we hypothesized that in an AN group, 10 years after onset of eating disorder, there would be (1) a decline in general cognitive functioning over time according to the WAIS-R compared with three years earlier, (2) deficits in intelligence level, cognitive profile (including a trough on the subtest of Object Assembly), executive functions and memory compared with matched healthy controls, (3) a test profile similar to that of children and adults diagnosed with autistic disorder in the AN subgroup with ASD.

SUBJECTS AND METHODS

Subjects

One hundred and two individuals, 51 who had DSM-III-R AN (APA, 1987) in the diagnostic study (Råstam, 1992) and their 51 comparison (COMP) cases, were invited to participate in the present research. The AN group comprised a total population sample of 24 individuals plus a mixed community-screening and referral group of 27 individuals (Råstam, 1992). The original diagnostic study was performed when these 51 individuals were on average 16 years old. The population group and the mixed group were similar in most respects analyzed (except with regard to consultation of services). Both these groups and a group of 51 sex-, age-, and school-matched comparison cases (originally selected by the school nurse) were examined psychiatrically at a mean age of 16 years using standardized measures.

The groups were re-examined on average 5 years (6 years after AN onset, Gillberg et al., 1996) and 8 years later, 10 years after AN onset (findings reported here), both by the psychiatrist of the diagnostic study,
and a psychiatrist/psychologist (and an additional psychiatrist in the second follow-up study blind to original group status), again using standardized measures. Of the 102 individuals (AN and COMP) in the diagnostic study, 47 AN and 51 COMP cases completed at least one of the neuropsychological tests (see below) at examination 8 years after the first study (10 years after mean reported onset of AN).

All 51 AN cases in this study had met DSM-III-R (APA, 1987) and DSM-IV (APA, 1994) criteria for the disorder. Forty-eight of these met or had met such criteria already at the first diagnostic study, but 3 were classified as partial at that time. However, these 3 cases later met full DSM-III-R and DSM-IV criteria.

The weight and height characteristics of the AN and COMP groups at ages 16, 21 and 24 years are shown in Table 1.

All 102 individuals, except three women in the AN group (who were interviewed over the telephone), were seen personally and examined in accordance with the procedure described below. Three individuals in the AN group still met criteria for AN, and another eleven had other eating disorder diagnoses (2 bulimia nervosa, 9 eating disorder not otherwise specified). One of the three telephone interviewees had persisting AN, thus, only two of those with current AN participated in the neuropsychological testing in the present study.

**Age of onset**

The mean age of AN reported onset was 14.3 years (95% confidence interval (CI): 13.9–14.7).

**Age at earlier examinations (study 1 and study 2)**

The mean age at first examination was 16.1 years (95% CI: 15.7–16.5) in the AN group and 16.0 years (95% CI: 15.5–16.5) in the COMP group. The mean age at 6-year follow-up was 21.0 years (95% CI: 20.5–21.4) in the AN group and 20.8 years (95% CI: 20.3–21.3) in the COMP group.

**Age at present follow-up study (study 3)**

The mean age at the present follow-up examination was 24.5 years (95% CI: 24.0–25.0) in the AN group and 24.2 years (95% CI: 23.7–24.7) in the COMP group.

**Follow-up period**

The average time that had elapsed from onset of AN to the time of the present study was 10.2 years (95% CI: 9.7–10.6). The average time that had elapsed from the first examination to the time of the follow-up study was 8.4 years (95% CI: 8.1–8.8) in the AN group and 8.1 years (95% CI: 7.7–8.4) in the COMP group.

**Methods used in study 1**

The individuals (AN and COMP) were seen and examined in great detail at the time of diagnosis of AN (at a mean age of 16 years) by a psychiatrist (MR). The instruments used have been reported on in Råstam (1992). No cognitive measures were used in study 1.

**TABLE 1**

Mean (SD) weight, height and BMI of AN and COMP groups at 16, 21, and 24 years

<table>
<thead>
<tr>
<th></th>
<th>AN (n = 51) Age (y)</th>
<th>COMP (n = 51) Age (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16a</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>49.4 (8.8)***</td>
<td>58.9 (11.8)</td>
</tr>
<tr>
<td>95 % CI</td>
<td>47.0–51.8</td>
<td>55.6–62.2</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>164.3 (5.8)</td>
<td>166.2 (6.4)*</td>
</tr>
<tr>
<td>95 % CI</td>
<td>162.7–165.9</td>
<td>164.4–168.0</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>18.3 (2.9)</td>
<td>21.2 (3.5)</td>
</tr>
<tr>
<td>95 % CI</td>
<td>17.5–19.1</td>
<td>20.2–22.2</td>
</tr>
</tbody>
</table>

*Note. AN: anorexia nervosa group, COMP: comparison group, CI: confidence interval.
*aAverage minimum BMI 14.9 kg/m² (SD 2.6) in group with AN (Råstam, 1992).
*p < .05 AN versus COMP at age 21 years.
**p < .01 AN versus COMP at age 16 years.

At 24 years mean weight was based on 45 individuals with AN and 49 COMP, mean height was based on 47 individuals with AN and 51 COMP, and mean BMI was based on 45 individuals with AN and 49 COMP. Note: Comparisons across groups, individuals with AN versus COMP, were only performed on one study at a time.
Methods used in study 2

All individuals were again examined by the psychiatrist who had performed the diagnostic study (MR), and by another psychiatrist /psychologist (ICG) blind to the original group status. Both doctors made independent ratings of each individual’s capacity of empathising with the perspectives, thoughts and feelings of others. The instruments used have been described in Gillberg et al. (1994). The WAIS-R was administered by ICG in all cases in this study. No other neuropsychological tests were used in study 2.

Methods used in study 3 (present study)

A number of different instruments were used in order to check on diagnostic stability over time (e.g., the SCID-I and SCID-II, on the basis of which DSM-III-R and DSM-IV diagnoses of axis I and axis II disorders were made) and to assess overall outcome status (e.g., the Morgan-Russell scales and the Global Assessment Functioning Scale, GAF (APA, 1994)). The instruments used and design of the overall study is described in detail in Wentz, Gillberg, Gillberg, & Råstam, 2001. Twenty AN and 14 COMP cases were assigned a psychiatric Axis I diagnosis (APA, 1994). Affective disorders occurred in 5 individuals in the AN group and 2 individuals in the COMP group. The mean GAF scores were significantly lower in the AN group, 65.3 (95% CI: 61.0 to 69.7) compared to 84.8 (95% CI: 81.7 to 87.9) in the COMP group (p < .0001). For the purpose of the present study only the neuropsychological test methods will be presented in more detail.

Neuropsychological tests used at 24 years

**WAIS-R**

ICG administered the WAIS-R (Wechsler, 1990). The version used was the recent Swedish translation (Bartfai, Nyman, & Stegman, 1992), which has a mean population IQ of 100, a verbal and a performance part, both with a population mean of 100, and 11 subtests yielding scaled score results of 1–19 (with a mean of 10). Forty-seven AN and 51 COMP individuals were tested with the WAIS-R.

**Wisconsin Card Sorting Test (WCST)**

The computerized version of the Wisconsin Card Sorting Test (WCST) (Berg, 1948; Heaton, 1981) was used. The WCST consists of four stimulus cards and 128 response cards that depict figure of varying forms or shapes (crosses, circles, triangles, or stars), colors (red, blue, yellow, or green) and numbers of figures (one, two, three, or four). The WCST is considered to be a measure of abstract reasoning ability and the ability to shift cognitive set. It is generally regarded as a test of “executive functions”, requiring the ability to maintain an appropriate problem-solving strategy across changing stimulus conditions in order to achieve a future goal. Forty-one AN and 47 COMP individuals completed the WCST.

**The Luria memory ten-word retrieval test**

Participants were read a standardised list of 10 words and were then asked to name as many of these words as possible immediately after having heard them and an hour later (Luria, 1966). Forty-seven AN and 51 COMP individuals were given this Luria ten word test.

**Test of facial recognition**

This test examines the ability to recognise faces without involving a memory component. The participants were instructed to match identical front views, front with side views and front views taken under different lighting conditions (Benton, Hamsher, Varney, & Spreen, 1983). The test was completed by 47 and 51 COMP subjects.

**The Birmingham Object Recognition Battery (BORB)**

We used subtests from the BORB, which contains a set of standardized procedures for assessing Neuropsychological disorders of visual object recognition: (a) the Minimal Feature View Test, and (b) the Foreshortened View Task (Riddoch & Humphreys, 1993). The participants were asked to match objects that differed in viewpoint. The BORB was administered to 40 AN and 47 COMP individuals.

**Statistical methods used**

The Wilcoxon (Mann-Whitney) rank sum test was used when comparing data at 6-year follow-up and 10-year follow-up and when comparing data across groups at 10-year follow-up. The cut-off level (=critical alpha) for statistical significance was set at 0.05. A possible relationship between the FSIQ and GAF was evaluated by the Spearman rank order correlation coefficient. Holm’s stepwise correction (Holm, 1979) was applied to adjust the p-values for the large number of statistical tests.
performed. In respect of our main hypotheses, Holm’s stepwise correction was not considered appropriate, since differences, if significant, would not be considered chance occurrences.

RESULTS

WAIS-R full scale IQ

Mean full scale WAIS-R IQ (FSIQ) in the AN group was not far from the mean of the general population, neither at age 21 or 24 years. Mean FSIQ in the AN group was 105.2 (95% CI: 101.3–109.0) at age 24 years, a significantly higher level than at age 21 years when it was 103.2 (p < .001). FSIQ in the COMP group had also increased from a mean of 106.5 to 109.4 (95% CI: 106.6–112.2) (p < .01). No individual scored under IQ 70 in any of the two subject groups. However, the AN group performed significantly lower on the FSIQ at age 24 years compared with the COMP group (p < .05) (Table 2).

WAIS-R Verbal versus performance IQ

Verbal IQ (VIQ) was lower than performance IQ (PIQ) in both groups (p < .001). PIQ — but not VIQ — increased significantly in both groups over time (AN: p < .001, COMP: p < .01).

WAIS-R subtest results

The AN group scored lower than the COMP group on the Object Assembly subtest (p < .02) (Table 2).

WAIS-R test profiles in the AN group with autism spectrum disorders (ASD)

The AN group diagnosed with ASD (Asperger syndrome or autistic-like condition) on both follow-up occasions and in connection with the first diagnostic assessment (=AN with ASD) (n = 8, 7 of whom were given the WAIS-R) scored significantly lower on the Coding subscale compared with the remaining cases with AN (n = 40) and the COMP group (AN with ASD: 7.9, 95% CI: 4.6–11.1, AN without ASD: 11.1, 95% CI: 10.3–11.9, p < .01; COMP:11.0, 95% CI: 10.3–11.7, p < .01) After Holm’s stepwise correction this finding failed to reach statistical significance. (Figure 1).

The Wisconsin Card Sorting Test (WCST)

There were no significant differences across the two groups on any of the WCST subtests (number of categories: AN: mean: 5.0, SD: 1.9; COMP: mean: 5.6, SD: 1.0; perseverative errors: AN: mean: 14.8, SD: 10.5; COMP: mean: 11.7, SD: 7.5; failure to maintain set: AN: mean: 1.2, SD: 1.3; COMP: mean: 1.1, SD: 1.1; incorrect

| Table 2 |

| IQ (FSIQ, VIQ, PIQ and subtest scaled scores) in AN and COMP groups |
|-------------------------------------------------|------------------|------------------|
| **WAIS-R IQ Test/Subtest** | **AN (n = 47)** | **COMP (n = 51)** |
| | 95% CI | Mean | 95% CI | Mean |
| FSIQ | 101.3–109.0 | 105.2* | 106.6–112.2 | 109.4 |
| VIQ | 97.5–104.6 | 101.0** | 101.5–106.6 | 104.0 |
| PIQ | 105.7–113.9 | 109.8 | 111.1–118.0 | 114.6 |
| Information | 10.1–11.4 | 10.7 | 10.9–12.0 | 11.5 |
| Digit span | 8.5–10.0 | 9.2 | 9.2–10.6 | 9.9 |
| Vocabulary | 8.7–9.9 | 9.3 | 9.3–10.2 | 9.8 |
| Arithmetic | 9.8–11.5 | 10.7 | 10.6–12.1 | 11.4 |
| Comprehension | 9.6–10.9 | 10.3 | 9.8–10.9 | 10.3 |
| Similarities | 9.9–11.2 | 10.5 | 10.7–12.0 | 11.3 |
| Picture completion | 11.0–12.4 | 11.7 | 10.8–12.1 | 11.5 |
| Picture arrangement | 10.6–12.3 | 11.5 | 11.3–12.7 | 12.0 |
| Block design | 11.2–12.9 | 12.1 | 12.1–13.8 | 13.0 |
| Object assembly | 10.5–11.9 | 11.2*** | 11.7–13.3 | 12.5 |
| Coding | 9.8–11.4 | 10.6 | 10.3–11.7 | 11.0 |

Note: CI: confidence interval, AN: anorexia nervosa group, COMP: comparison group. FSIQ: full scale IQ, VIQ: verbal IQ, PIQ: performance IQ.

*p < .05 AN versus COMP group.

**p < .001 VIQ versus PIQ in AN and COMP group.

***p < .02 AN versus COMP group.
responses: AN: mean: 29.6, SD: 21.1; COMP: mean: 23.0, SD: 15.1).

The Luria memory test (ten word retrieval)

There were no significant differences across the two groups on the Luria memory test (Table 3).

The Facial Recognition Test

There were no differences across the two subject groups on the Facial Recognition Test (Test 1: mean duration (seconds): AN: 29.0, SD: 16.0; COMP: 29.0, SD: 15.6; Test 1: mean errors: AN: 0, COMP: 0.04, SD: 0.20; Test 2: mean duration (seconds): AN: 145.3, SD: 104.4; COMP: 137.5, SD: 67.0; Test 2: mean errors: AN: 2.6, SD: 2.0; COMP: 2.4, SD: 1.3).

The Birmingham Object Recognition Battery (BORB)

There were no differences across the subject groups on the BORB (Test 1: AN: mean: 24.9, SD: 0.3; COMP: mean: 24.9, SD: 0.3; Test 2: AN: mean: 24.5, SD: 0.5; COMP: mean: 24.4, SD: 0.6).

IQ related to some variables

There was a positive relationship between FSIQ and social functioning according to the GAF ($p < .01$). FSIQ did not differ between the subgroup with current axis I disorder (other than ED) and those with no axis I disorder, or, in the AN group, between those with persisting ED from those without persisting ED.

DISCUSSION

This controlled neuropsychological follow-up of AN 10 years after teenage onset (and several years after refeeding in almost 95% of AN cases) confirms the impression from the follow-up 6 years after AN-onset that overall IQ is in the normal range in the vast majority of cases. Both groups had improved their FSIQ since the previous study. However, FSIQ was now significantly higher in the COMP group. In keeping with the finding almost 4 years previously, the AN group exhibited poorer performance on the Object Assembly subtest. The significant minority with persistent ASD (about 15%) in the AN group continued to show a neuropsychological test pattern at least partly consistent with that shown by other high-functioning individuals in the autism spectrum. Executive function deficits and major memory problems were rare both in the AN and comparison groups.

Our findings are important because they are likely to be representative of all teenage onset cases with AN. We are not aware of any other population-based study of neuropsychological results from a population-representative group of individuals with AN. They underscore a number of clinically important tentative conclusions:
First, AN is not associated with high IQ as has been suggested on the basis of results obtained in selected clinical samples (Blanz et al., 1997). In clinical practice, this means that individuals with AN should not be overestimated as regards their level of intellectual functioning. Clearly, individual variation is great and individual assessment rather than generalized assumptions should guide clinical evaluation and interventions.

Second, in spite of verbal IQ being lower than performance IQ in AN, this should not be taken as a general indication of poor verbal performance in AN, given that the pattern was similar in the COMP group.

Third, the existence of a considerable minority of ASD cases in a representative sample of individuals with AN was confirmed, and the test profile of that group reified. But results concerning WISC/WAIS profile are not consistent. Most studies show that individuals with high functioning autism and Asperger syndrome have a profile characterized by a high score on Block Design with a low Comprehension score (Gilchrist et al., 2001; Goldstein, Beers, Siegel, & Minshew, 2001; Goldstein, Minshew, Allen, & Seaton, 2002). One study found a peak on Block Design in children with autism (Ehlers et al., 1997), while an Asperger syndrome group had significantly better verbal ability than the autism group and troughs on Object Assembly and Coding.

The tendency for poor results on the Coding subtest of the Wechsler scales has been repeatedly shown to apply in Asperger syndrome (Ehlers, 1997; Gillberg & Cederlund, 2004), one of the most common of the ASD. Even though the difference across the AN subgroups with and without ASD was no longer significant after Holm’s correction, we believe that a true difference exists which is obscured by the limited number of cases in the ASD subgroup. The relatively poor result on this subtest could be taken to mirror the extreme slowness, circumstance and drive for perfection often found in Asperger syndrome (Ehlers, 1997). Alternatively, it could be seen to reflect parietal dysfunction which has been shown to be the effect of starvation on the brain and which has been demonstrated to be quite prevalent in post-starvation AN (Frank et al., 2002; Gruwalnd et al., 2001; Hendren, De Backer, & Pandina, 2000, Råstam et al., 2001). We have argued that the communicative style with a severe empathy deficit is also indirectly supportive of a similarity between AN and ASD (Råstam, Gillberg, & Johansson, 1997).

The test we used to tap executive dysfunction may not have been sensitive enough to reveal such problems in AN. Verbal memory appeared unaffected in AN, at least as judged on the basis of results on the Luria test. More subtle memory deficits would not have been picked up by this rather crude assessment of mnemonic function.

In conclusion, our study provides no support for the notion of high FSIQ in AN. Our hypothesis that FSIQ would decline over time in the AN group was not supported. However, the gap between the AN and the COMP group (in favour of the COMP group) appeared to have widened compared to test results at age 21 years, and so an overall effect on general cognitive functioning of teenage onset AN cannot be ruled out. As hypothesized, the object assembly subtest of the WAIS-R still discriminated between the AN and COMP group. Executive functions and verbal memory appeared to be relatively unimpaired in the longer-term perspective of AN after refeeding.

Original manuscript received 5 August 2005
Revised manuscript accepted 6 January 2006
First published online 20 October 2006

REFERENCES


Uher, R., Murphy, T., Friederich, H. C., Dalgleish, T., Brammer, M. J., Giampietro, V., Phillips, M. L., Andrew, C. M., Ng, V. W., Williams, S. C., Campbell, I. C., & Treasure, J. (2005). Functional neuroanatomy...


