Understanding cognitive heterogeneity in psychosis and high risk individuals
Meijer, J.H.

Citation for published version (APA):
Meijer, J. H. (2012). Understanding cognitive heterogeneity in psychosis and high risk individuals

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
COGNITIVE ALTERATIONS IN PATIENTS WITH NON-AFFECTIVE PSYCHOTIC DISORDER AND THEIR UNAFFECTED SIBLINGS AND PARENTS

Julia H. Meijer*, Claudia J.P. Simons*, Piotr J. Quee, Kim Verweij, GROUP Investigators* # joint first authors

# GROUP investigators: René S. Kahn, Don H. Linszen, Jim van Os, Durk Wiersma, Richard Bruggeman, Wiepke Cahn, Lieuwe de Haan, Lydia Krabbendam and Inez Myin-Germey

ABSTRACT

Objective: The purpose of this study was to examine a range of cognitive measures as candidate phenotypic liability markers for psychosis in a uniquely large sample of patients with psychosis, their unaffected relatives and control subjects.

Method: Patients with non-affective psychosis (n=1093), their unaffected siblings (n=1044), parents (n=911), and controls (n=587) completed a comprehensive cognitive test battery. Cognitive functioning was compared using tests of verbal learning and memory, attention/vigilance, working memory, processing speed, reasoning and problem solving, acquired knowledge and social cognition. Age- and gender-adjusted z-scores were compared between groups using mixed-model analyses of covariance. Clinically relevant impairment (-1 and -2 SD from control mean) was compared between subject groups.

Results: Patients performed significantly worse than controls on all cognitive domains (z-range -0.26 to -1.34). Siblings and parents showed alterations for immediate verbal learning, processing speed, reasoning and problem solving, acquired knowledge and working memory (z-range -0.22 to -0.98). Parents showed additional alterations for social cognition. Prevalence of clinically relevant impairment in relatives ranged from 50% (-1 SD criterion) to 10% (-2 SD criterion).

Conclusion: Unaffected family members show mild alterations in specific cognitive domains that can be further tested as intermediate phenotypes in genetically sensitive analyses.

Significant outcomes:
» Patients with non-affective psychotic disorder are characterized by cognitive alterations across all cognitive domains with relatively small effect sizes compared to meta-analytic results.
» Verbal learning, processing speed, reasoning and problem solving, working memory, and acquired knowledge are the most promising cognitive intermediate phenotypes, demonstrating alterations in genetic high risk groups.
» The distribution of clinically relevant cognitive alterations in patients and their first-degree relatives is suggestive of a continuum of neuropsychological functioning, with approximately 30% of the patients and 50% of the relatives displaying no clinically manifest deficit.

Limitations:
» Effect sizes in parents might have been somewhat inflated due to differences in age-range between the parent and control group.
» Not all subjects had complete cognitive test scores, which may have impacted on the effect sizes.
» Group differences in educational achievement and IQ remain a potential explanation for group differences on other cognitive test scores and cannot be ruled out through statistical adjustment.
INTRODUCTION

Cognitive alteration is a stable, trait-related aspect of schizophrenia that has been associated with impaired quality of life and poorer functional outcome (1). Subtle cognitive deficits are present prior to psychosis onset and may help to predict conversion to psychosis in “clinical high risk” subjects who are in the putative prodromal phase of psychosis (2, 3). Attenuated cognitive alterations have also been reported in clinically unaffected relatives of schizophrenia patients who are referred to as being at “genetic high risk” for psychosis. Cognitive alterations may therefore reflect the expression of genetic vulnerability for schizophrenia (4-7). Identifying such cognitive intermediate phenotypes may be a productive approach in genetic linkage and association studies in schizophrenia as they are closer to the mechanism for gene action than the overall disease phenotype.

Although evidence on cognitive alterations as intermediate phenotypes in schizophrenia is promising, sample sizes have been limited (8). This is illustrated by the fact that the most recent review in young relatives of psychotic patients included 18 studies with a mean of 102 high-risk relatives (range 29-322) and 84 control subjects (range 26-201), while only 5 studies included a patient group (mean 76, range 27-207) (4). Thus, the most appealing evidence to date originates from meta-analyses and reviews. Summarizing evidence related to a specific hypothesis can be distorted however by the selective publication of studies with certain (especially positive) results (9). In addition, studies on cognition in genetic high risk samples have been limited by (i) the analysis of different biological relatives as one group (siblings, children and parents), (ii) the inappropriate screening for psychiatric disorders in the relatives, and (iii) limited assessment of cognitive functions (8). In combination with a high within- and between-subject variability of cognitive performance in genetic high risk samples (8), these methodological limitations have hampered the identification of cognitive intermediate phenotypes so far. There is also a lack of knowledge about the proportion of genetic high risk subjects that are cognitively indistinguishable from healthy controls and which proportion displays a clinically relevant cognitive impairment. While for patients with schizophrenia the proportion of cognitively spared patients lies around 15-30% (8, 10), little is known about this percentage in genetic high risk subjects.

AIMS OF THE STUDY

The aim of the present study was to test a broad range of cognitive measures as candidate intermediate phenotypes in a large population of patients with a non-affective psychotic disorder, their unaffected siblings and parents, and control subjects. Therefore, age- and gender-adjusted z-scores were compared between subject groups. Secondly, the proportions of clinically relevant cognitive impairment (no, mild, moderate, severe) were compared between subject groups, using both 1 and 2 SD below the control mean as impairment cut-off.
MATERIAL AND METHODS

The sample of the Genetic Risk and Outcome of Psychosis study (GROUP) was described previously (11). The baseline GROUP sample consists of 1120 patients with a non-affective psychotic disorder, 1057 unaffected siblings, 919 parents, and 590 unrelated control subjects. Within the patient group, 84% had a schizophrenia-related disorder (DSM-IV-TR code 295.x, 80.0% schizophrenia, 13.1% schizo-affective disorder, 6.9% schizophreniform disorder; n=945), 1% was diagnosed with psychotic illness in the context of substance abuse or somatic illness (DSM-IV-TR code 293.x/292.x, n=9), and 13% fulfilled criteria for other psychotic disorders (DSM-IV-TR code 297/298, n=149). Six patients had a missing diagnosis but fulfilled inclusion criteria, and 11 patients had a final diagnosis of affective psychosis after fulfilling diagnostic criteria of non-affective psychosis at study entry, which may have been the result of subtle diagnostic changes between the time of identification for inclusion and actual assessment.

The mean age at onset of psychosis was 22.6±6.7 years (range 10-54), and the mean illness duration was 4.4±4.1, years (range 0.1-41.1). At the time of testing 86.1% of the patients were on antipsychotic treatment with one or more antipsychotics, with olanzapine (27.8%), risperidon (23.7%), and clozapine (11.6%) being the antipsychotics most frequently used. The mean number of psychotic episodes was 1.7±1.1 (range 1-8) and the mean number of psychiatric hospitalizations was 1.9±2.3 (range 0-30). According to PANSS remission criteria (12, 13), 45.1% of the patients were in remission from psychosis at the time of testing. Patients had a mean PANSS positive score of 14.0±6.4 (range 7-41) and a mean PANSS negative score of 15.0±6.6 (range 7-39). In 12.1% of the siblings (n=152), 19.4% of the parents (n=178) and 10.0% of the controls (n=59) there was a lifetime history of a DSM-IV-TR mental disorder. Depressive disorders were by far most common, reported in 10.5% of siblings (n=111), in 17.7% of parents (n=163) and in 8.5% of controls (n=50).

Verbal learning and memory was assessed with a visually-presented Word Learning Task (WLT; 14). Outcome measures were: immediate recall (number of words recalled over the three 15-word trials), retention rate (delayed free recall after 20 minutes divided by the maximum score of immediate recall trials 1-3), and recognition (true positives – false positives). Attention/vigilance was assessed using a Continuous Performance Test (CPT-HQ) with working memory load, known in the literature as CPT-AX (15). Outcome measures were reaction time (reaction time for correct detections) and accuracy (proportion of correct detections). The WAIS-III subtest Arithmetic (16) was assessed to measure working memory. This subtest is a relatively complex measure of working memory capacity since it also addresses verbal comprehension and arithmetic skills, which are both associated with educational level (17). The Response Shifting Task (RST), a modified version of the Competing Programs Task (18, 19), was administered in order to assess set-shifting ability from an imitation response rule to a reversal response rule. Outcome variables were accuracy cost (proportion correct in the imitation condition – proportion correct in the reversal condition) and reaction time cost (reaction time in the reversal condition – reaction time in the imitation condition). The first response in each block and responses that were preceded by errors were excluded from analyses (20). In addition, only reaction times for correct responses were used and trials with a reaction time shorter than 150ms were eliminated from the analyses. The WAIS-III subtest Digit Symbol-Coding (16) was
used as a measure of speed of processing. Reasoning and problem solving was assessed using the subtest **Block Design** from the WAIS-III (16). The WAIS-III **Information** subtest (16) was used as a measure of acquired knowledge. Moreover, the cognitive assessment included two dimensions of social cognition. Outcome measures of the **Degraded Facial Affect Recognition** task (DFAR; 21) were the proportion correctly recognized neutral, happy, fearful and angry faces and the overall proportion correct. The short form of the **Benton Facial Recognition Test** (22) was assessed to be used as a covariate to adjust for non-emotional facial processing skills. Theory of mind was assessed using the **Hinting Task** that assesses mentalizing capacity required to comprehend real intentions behind indirect speech (23). Outcome measure was the sum of the ten separate item scores (range 0-20). It took approximately 90 to 120 minutes to complete the neuropsychological tests that were administered in the following fixed order: WLT immediate recall, RST, CPT-HQ, Digit Symbol-Coding, WLT delayed recall, WLT recognition, DFAR, BFRT, Information, Arithmetic, Block Design, Hinting Task.

**Statistical analyses**

The data were analyzed using the SPSS 17.0 statistical package. To facilitate the comparison of cognitive functioning between patients, siblings, parents and control subjects, raw scores were converted into z-scores. Due to the fact that patients and siblings belong to different age categories than parents (Table 1), z-scores were adjusted for age and gender by dividing the control group into reference groups, setting a minimum of 50 subjects per stratum. This resulted in eight categories following methods described by Keefe et al. (24):

- Age ≤20: 71 males and 64 females
- Age 21-30: 85 males and 105 females
- Age 31-40: 59 males and 65 females
- Age ≥41: 55 males and 85 females

Adjusted z-scores were computed as follows. Let $X_{jk}$ be the raw score $X$ on subtest $j$ ($j=1$ to 15) for subject $k$. Assume that subject $k$ has sex $l$ ($1 =$ male and 2 = female) and is in age category $m$ ($1$≤20, 2=21-30, 3=31-40, 4=41 years). The scaled score is then computed as follows: $z_{corrected} = (X_{jk} - M_{jlm}) / SD_{jlm}$; where $M_{jlm}$ and $SD_{jlm}$ are the mean and standard deviation, respectively, for test $j$ of the control population for sex $l$ and age category $m$. Resulting z-scores are identical to Glass’s delta estimator of effect size (25). Observations with more than three SDs from the mean were considered outliers and were replaced by the mean plus or minus 3 times the SD.

Subsequently, adjusted z-scores were compared between patients/siblings/parents and control subjects. In addition, adjusted z-scores were compared between siblings/parents and patients. To control for intra-family correlation mixed model analyses of covariance (ANCOVAs) were performed in which family was used as a random factor with a random intercept. Status (patient, sibling, parent, control) was the independent variable. Dependent variables were adjusted z-scores for 18 outcome measures derived from 10 cognitive tests. Although educational level of the subject and IQ may be associated with many of the putative intermediate phenotypes in question, they are also powerfully affected by schizophrenia (26). Therefore, the highest educational degree that had been obtained by one of the parents was entered into the model as a covariate instead. Since the Dutch educational system already
differentiates after primary school, we have chosen a coding system other than years of education. This ordinal 8-point scale indicates the level of education, and ranges from primary school to university. Mixed-model ANCOVAs for the DFAR variables incorporated the BFRT test scores as an additional covariate. Since mixed-model ANCOVAs were performed for multiple cognitive outcome parameters (n=18), a Bonferroni correction was adopted by setting the alpha level to .05/18 = .0028.

Subsequently, between-group comparisons were performed only in those cognitive parameters for which the effect of Status in the ANCOVA was significant. Five post-hoc analyses were performed for each of those cognitive parameters, firstly to compare patients/siblings/parents with controls, and secondly to compare siblings/parents with patients. For these post-hoc analyses the alpha value was set to .05 / (5 x the number of cognitive parameters for which post-hoc comparisons were performed). In addition, the same correction for multiple analyses was also applied for the much more conservative alpha value of .001.

Normality of cognitive outcome measures was visually inspected and confirmed if the test-statistic W in the Shapiro Wilk test exceeded 0.90. All but three outcome measures were normally distributed. Ceiling effects were present for CPT accuracy, WLT recognition and the Hinting Task. Data transformation did not improve the normality of the distributions; therefore both parametric and non-parametric testing was conducted. Secondary to the mixed-model ANCOVAs, group-comparisons were performed using Kruskal-Wallis tests that were followed by post-hoc Mann-Whitney tests with Bonferroni correction. To minimize the risk of type-I errors, the analyses that yielded the most conservative results were chosen for further discussion.

Secondly, raw test scores were converted into dichotomous variables of ‘impaired’ or ‘not impaired’. A priori, both 1 SD (27-29) and 2 SD (30) below age- and gender-corrected control mean were selected as cut-off for clinical impairment. For cognitive tests with more than one outcome measure (e.g. RST reaction time and RST accuracy), an impairment was deemed present if the score of at least one measure was below the cut-off. Impairment scores were summed to generate total impairment scores (range 0-10). Based on the control mean of 1.8 tests with an impairment (based on 1 SD cutt-off), the criterion for ‘not impaired’ was defined as 0-2 tests with an impairment. For the width of the following categories, the control SD of 1.5 was used, resulting in the categories ‘mild impairment’ (impairment on 3-4 tests), ‘moderate impairment’ (impairment on 5-6 tests), or ‘severe impairment’ (impairment on 7 or more tests). Total impairment scores were calculated for subjects that had completed at least 9 out of 10 cognitive tests. Chi-square tests were used to detect statistically significant differences in total impairment scores between the subject groups. A Bonferroni correction was applied by setting the alpha value to .05/16 = 0.003, since 4 dependent variables were compared between 4 subject groups. Analyses were performed using SPSS 17.0 for Windows. Release 2.0 of the GROUP database was used for the analyses.
RESULTS

Data inspection

Although WLT recognition was assessed for 89.3% of the subjects, for only 47.7% of all subjects reliable data were available for this task. This was mainly due to technical problems. These test scores were included into the mixed-model ANCOVAs, but not in the calculation of total impairment scores that allowed for a maximum of one missing value. Inspection of the missing values for the remaining nine cognitive tests showed that 2922 subjects (79.3%) had completed all tests, while 453 individuals (12.3%) had missing data for one test and 260 individuals (7.1%) had missing data for more than three tests. For 49 individuals (1.3%) no cognitive test results were obtained, so these subjects were excluded from the analyses. The mean proportion of missing tests was 6.0% (range 3.4%-11.4%). Test duration and test rank were not associated with proportion of missings, but tasks with computerized scoring had a higher mean proportion of missings (9.9%) than the paper-pencil scoring tasks (4.1%). One particularly interesting finding is the relatively low proportion of missing values (4.0%) for WLT immediate recall, for which stimuli were presented in a computerized fashion, but with paper-pencil scoring. When comparing demographic variables of subjects categorized by the number of tests missing (no missings, 1-3 tests missing, >3 missings, not tested), no significant differences were found for gender, highest educational degree, or age. Patient status however was associated with the number of missing values. The proportion of missings between patients and the other three subject groups taken together (n=2565) was 12.5% vs. 12.2% (1-3 tests missing), 9.5% vs. 6.0% (>3 tests missing) and 2.4% vs. 0.9% (not tested), \( \chi^2(3) = 28.27, \ P<0.001 \).

Group comparisons

Patients were significantly more often male compared to siblings, parents, and controls. The mean age of patients and siblings was lower than the mean age in controls and in parents. Furthermore, there were statistical differences between the four subject groups in IQ, educational degree of the subject, and educational degree of the parent (table 1).

Observed means and standard deviations for cognitive test scores and results from mixed-model ANCOVAs are presented in table 2. Since mixed-model ANCOVAs were significant for 15 out of 18 cognitive outcome measures, post-hoc tests were denoted as significant at the 0.05 level if the \( P \)-value was smaller than 0.0007, resulting from 0.05/(5x15). Moreover, post-hoc tests were denoted as significant at the 0.001 level if the \( P \)-value was smaller than 0.00001, resulting from 0.001/(5x15). Age- and gender-adjusted z-scores in patients, siblings, and parents are displayed in figure 1. Because higher scores reflected worse performance for CPT-HQ reaction time, RST reaction time cost, and RST accuracy cost, z-scores for these measures were inverted.

For the three non-normally distributed tests, non-parametric testing yielded somewhat different results compared to results from mixed-model ANCOVAs. Mann-Whitney testing yielded significantly worse performance on the CPT accuracy in parents (Md(n)=583.65) compared to control subjects (Md(n)=815.42), U=143156.50, Z=-10.71, \( P<0.001, r=-0.29 \). In addition, Mann-Whitney testing did not yield significant results on the WLT recognition for the comparison between patients (Md(n)=451.21) and parents (Md(n)=499.78), U=99099.00, Z=-2.72,
$P_{\gamma} = 0.450, r=-0.09$, and for the comparison between siblings ($Mdn=386.20$) and control subjects ($Mdn=442.00$), $U=66659.50, Z=-3.28, P=0.075, r=0.12$. For the Hinting Task, no differences emerged between parametric and non-parametric testing. For WLT recognition, the more conservative non-parametric results were chosen over results from mixed-model ANCOVAs. For the other two tasks, the parametric results were maintained.

Against the background of recent findings (31) the analyses were repeated with cannabis use (current, lifetime, or never) as an additional covariate. Co-varying for cannabis, however, did not change any of the group-comparisons from significant to non-significant or vice versa (results not shown).

**Total impairment scores**

Figure 2 shows that with a cut-off of 1 SD below control mean, 29.6% of patients are classified as having no cognitive impairment against 71.4% of controls, with in-between rates for parents and siblings ($\chi^2(3)=214.4, P<0.05$). While the proportions of mild impairment did not differ between subject groups ($\chi^2(3)=10.4, P=0.12$), both moderate ($\chi^2(3)=86.3, P<0.05$) and severe impairment ($\chi^2(3)=127.6, P<0.05$) showed significant differences between groups. Patients showed the highest proportion of moderate and severe impairment and control subjects the lowest, with in-between rates for parents and siblings. Figure 3 shows that with a cut-off score of 2 SD, the rate of subjects classified as having no impairment increased to 70.8% in patients and 98.2% in controls, with in-between rates for parents and siblings ($\chi^2(3)=225.3, P<0.05$). Mild impairment rates decreased to 18.5% in patients and 1.3% in controls, $\chi^2(3)=114.3, P<0.05$. While 7.8% of patients displayed moderate impairment, this was very low in the other three subject groups ($\chi^2(3)=93.0, P<0.05$). Severe impairment was rare in patients (2.9%), very rare in parents and siblings (0.2% and 0.3%, respectively) and absent in controls, $\chi^2(3)=46.8, P<0.05$.

**DISCUSSION**

In a uniquely large sample of patients with non-affective psychotic disorder, their unaffected relatives and control subjects, this study has replicated and extended meta-analytic results

| Table 1. Socio-demographic characteristics of patients, siblings, parents and controls |
|----------------------------------------|----------------|----------------|----------------|----------------|
|                                      | Patients $n=1093$ | Siblings $n=1044$ | Parents $n=911$ | Controls $n=587$ |
| % Male gender                        | 76.2%           | 45.8%          | 42.8%          | 45.5%          |
| Age (years)                          | 27.7±8.1        | 27.8±8.3       | 54.8±6.9       | 30.4±10.6      |
| WAIS-III estimated IQ                | 94.9±16.1       | 102.6±15.5     | 103.1±17.0     | 109.6±15.2     |
| Educational degree subject          | 4.1±2.0         | 5.1±2.1        | 5.1±2.3        | 5.4±1.8        |
| Educational degree parent           | 5.2±2.4         | 5.2±2.4        | 3.4±2.3        | 5.0±2.4        |

*a* Wechsler-Adult Intelligence Scale short form (55)
Figure 1. Age- and gender-corrected z-scores for patients, siblings, parents and controls. WLT = word learning task; Immed = immediate recall; Retent = retention rate; RST = response shifting task; RST RT= reaction time cost; RST Acc= accuracy cost; CPT = continuous performance test; CPT RT= CPT reaction time; CPT Acc= CPT accuracy; Digit symb= digit symbol coding; Inform= information; Arithm= arithmetic; Block Des= block design; DFAR = degraded facial affect recognition; Neutr= neutral; BFRT = Benton facial recognition test; Hint= Hinting Task.

Figure 2a and 2b. Proportions of ‘no’ (0-2 tests altered), ‘mild’ (3-4 tests altered), ‘moderate’ (5-6 tests altered) and ‘severe’ alteration (7 or more tests altered) for each subject group. Cut-off scores for alteration are ≤-1 SD (figure 2a) and ≤-2 SD (figure 2b) from the control mean.
Table 2. Observed means and standard deviations of cognitive test scores and P-values of between-subject comparisons following mixed-model ANCOVAs using z-standardized scores and adjusting for parental education level.

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Patients (n=1093) M (SD)</th>
<th>Siblings (n=1044) M (SD)</th>
<th>Parents (n=911) M (SD)</th>
<th>Controls (n=587) M (SD)</th>
<th>Test Statistic (df)</th>
<th>P value Pat vs Co</th>
<th>P value Sib vs Co</th>
<th>P value Sib vs Pat</th>
<th>P value Par vs Co</th>
<th>P value Par vs Pat</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLT Immediate Recall</td>
<td>22.93 (6.09)</td>
<td>26.89 (5.77)</td>
<td>23.26 (6.11)</td>
<td>28.43 (5.38)</td>
<td>132.17 (3, 3111)</td>
<td>&lt;.001</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WLT Retention Rate</td>
<td>0.77 (0.21)</td>
<td>0.84 (0.17)</td>
<td>0.78 (0.20)</td>
<td>0.83 (0.16)</td>
<td>13.22 (3, 3105)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>-</td>
</tr>
<tr>
<td>WLT Recognition⁵</td>
<td>11.03 (3.48)</td>
<td>12.36 (2.99)</td>
<td>11.45 (3.33)</td>
<td>12.96 (2.13)</td>
<td>71.69 (3)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>RST RT Cost</td>
<td>205.84 (221.40)</td>
<td>198.43 (208.91)</td>
<td>216.42 (240.00)</td>
<td>194.78 (176.21)</td>
<td>0.72 (3, 2610)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RST Accuracy Cost</td>
<td>0.26 (0.27)</td>
<td>0.22 (0.25)</td>
<td>0.35 (0.33)</td>
<td>0.22 (0.25)</td>
<td>12.38 (3, 2841)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CPT Reaction Time</td>
<td>430.17 (84.43)</td>
<td>410.24 (78.39)</td>
<td>429.01 (81.90)</td>
<td>412.80 (82.67)</td>
<td>37.30 (3, 2856)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CPT Accuracy</td>
<td>98.75 (2.27)</td>
<td>99.51 (1.50)</td>
<td>98.89 (2.18)</td>
<td>99.63 (1.01)</td>
<td>40.62 (3, 2831)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Digit-Symbol coding</td>
<td>65.43 (16.26)</td>
<td>79.23 (15.44)</td>
<td>67.97 (16.72)</td>
<td>83.89 (14.60)</td>
<td>223.40 (3, 3108)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Information</td>
<td>16.78 (5.46)</td>
<td>16.83 (5.20)</td>
<td>17.61 (5.43)</td>
<td>18.82 (4.65)</td>
<td>33.69 (3, 3050)</td>
<td>&lt;.001</td>
<td>-</td>
<td>NS</td>
<td>NS</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

⁵: Significant at the .05 level.
<table>
<thead>
<tr>
<th>Task</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>F (df, df error)</th>
<th>p</th>
<th>p</th>
<th>p</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic</td>
<td>12.28 (4.78)</td>
<td>13.84 (4.43)</td>
<td>13.70 (4.25)</td>
<td>15.30 (4.16)</td>
<td>122.57 (3, 3110)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Block Design</td>
<td>40.47 (17.00)</td>
<td>44.87 (15.08)</td>
<td>32.15 (14.52)</td>
<td>46.55 (14.17)</td>
<td>80.30 (3, 3098)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.05</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DFAR Neutral</td>
<td>77.76 (17.75)</td>
<td>80.43 (15.03)</td>
<td>76.18 (16.95)</td>
<td>81.14 (15.17)</td>
<td>10.44 (3, 2981)</td>
<td>&lt;0.05</td>
<td>NS</td>
<td>&lt;0.05</td>
<td>NS</td>
</tr>
<tr>
<td>DFAR Happy</td>
<td>86.48 (13.07)</td>
<td>88.20 (10.72)</td>
<td>82.59 (13.66)</td>
<td>87.33 (11.12)</td>
<td>5.51 (3, 2981)</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>DFAR Fearful</td>
<td>47.23 (19.80)</td>
<td>52.54 (19.63)</td>
<td>48.35 (18.98)</td>
<td>53.75 (18.22)</td>
<td>23.39 (3, 2981)</td>
<td>&lt;0.05</td>
<td>NS</td>
<td>&lt;0.05</td>
<td>NS</td>
</tr>
<tr>
<td>DFAR Angry</td>
<td>62.12 (20.88)</td>
<td>68.81 (19.22)</td>
<td>60.61 (20.46)</td>
<td>70.43 (18.60)</td>
<td>30.06 (3, 2981)</td>
<td>&lt;0.05</td>
<td>NS</td>
<td>&lt;0.05</td>
<td>NS</td>
</tr>
<tr>
<td>DFAR total</td>
<td>68.40 (10.77)</td>
<td>72.50 (9.35)</td>
<td>66.93 (10.45)</td>
<td>73.16 (9.13)</td>
<td>40.78 (3, 2981)</td>
<td>&lt;0.05</td>
<td>NS</td>
<td>&lt;0.05</td>
<td>NS</td>
</tr>
<tr>
<td>BFRT</td>
<td>22.76 (2.31)</td>
<td>23.17 (2.16)</td>
<td>22.44 (2.44)</td>
<td>23.14 (2.05)</td>
<td>5.73 (3, 3178)</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Hinting Task</td>
<td>17.54 (2.78)</td>
<td>18.84 (1.66)</td>
<td>18.79 (1.62)</td>
<td>19.08 (1.31)</td>
<td>85.31 (3, 3101)</td>
<td>&lt;0.001</td>
<td>NS</td>
<td>&lt;0.001</td>
<td>NS</td>
</tr>
</tbody>
</table>

WLT = word learning test; RST = response shifting task; RT Cost = reaction time cost; CPT = continuous performance test; DFAR = degraded facial affect recognition; BFRT = Benton facial recognition test

*P-values after Bonferroni correction

*For WLT Recognition more conservative results from nonparametric analyses are presented
Unaffected siblings and parents performed intermediate between patients and control subjects on specific tasks. Alterations in verbal learning and memory, speed of processing, acquired knowledge, working memory, and reasoning and problem solving appear to be associated with a familial predisposition to psychotic disorder and may represent putative intermediate phenotypes with the potential of assisting in the genetic analyses of the psychosis phenotype. However, analyzing data in terms of cognitive alterations at the group level may obscure cognitive heterogeneity in the level of clinically relevant impairment. Especially given the relatively small effect sizes, it is important to explore how the mean cognitive alterations seen in patient and sibling groups translate into proportions of clinically relevant impairment.

For both criteria, cognitive impairments in relatives were more common compared to controls, but less common compared to patients. Based on the 1SD cut-off around 50% of the unaffected relatives displayed some level of cognitive impairment; however this proportion diminished to around 10% by adapting the 2SD cut-off for impairment. Results illustrate that although half of the unaffected relatives may experience some degree of cognitive difficulty, severe cognitive impairments are restricted to a small minority.

Results in patients correspond to meta-analytic evidence that the majority of psychotic patients display a generalized cognitive alteration extending across most cognitive domains, with effect sizes ranging from small to large and clinically relevant (32-35). The magnitude of severity of alterations in patients ranged from -0.18 to -1.34 SD from the control mean, with an average cognitive alteration in patients of -0.61 SD (-0.70 for neurocognition and -0.45 for social cognition). This is mild compared to the approximate average cognitive deficit of -1 SD suggested by previous research in schizophrenia patients (33). One possible explanation for these conservative findings is that setting an extreme of three standard deviations below the mean may have artificially truncated the true range of some cognitive tasks. Another possibility is that the higher percentage of missing data in the patient group may have selected out those who were too impaired to complete cognitive testing. Alternatively, the inclusion of tests that have been previously associated with premorbid intellectual functioning and education (Information, Arithmetic) into a composite score may have produced a measure that is not optimally representative of current neurocognitive impairment (17). Moreover, even though the majority of patients were diagnosed with schizophrenia-related disorders, the decision to include patients with other non-affective psychotic disorders may have attenuated effect sizes. However, differentiating between non-affective psychotic disorders is sometimes difficult, and focusing exclusively on the inclusion of schizophrenia patients may inflate effect sizes through selection bias. Despite the fact that cognitive dysfunction is not a DSM criterion for schizophrenia, psychotic patients who are cognitively intact may be more likely to be diagnosed with e.g. psychotic disorder NOS or substance-induced psychosis. Finally, it cannot be excluded that the prerequisite for patients to have siblings and/or parents able and willing to participate in the study may have selected out the more socially isolated and impaired patients.

The relatively small effect sizes are especially noteworthy given that education or IQ were not pursued as covariates. It may be argued that not controlling for these potential confounders may have inflated effect sizes. However, psychotic disorders are neurodevelopmental in nature,
and subtle deficits on neurocognitive measures during childhood and adolescence have been associated with an increased risk of non-affective psychosis (36). Controlling for education or IQ would thus be inappropriate given that they are powerfully affected by psychosis (26) and genetic vulnerability for psychosis (37). Adjusting for education or IQ would successfully ‘remove’ the variance due to education or IQ, but would not successfully ‘control for’ the variance due to education and IQ, which are meaningful components of the psychotic disorder phenotype (38). The present study therefore pursued parental education as covariate in the analyses instead, since parental education is less likely to be substantially related to psychotic disorder (38).

Traditional neuropsychological criteria classify individuals scoring -1 SD or more from the control mean as cognitively impaired (27-29). While patients’ scores on most cognitive outcome measures were significantly below the healthy control mean, only WLT immediate recall, CPT accuracy, Digit Symbol-Coding, Arithmetic and Hinting Task performance would be classified as impaired according to these criteria. This is in line with previous studies that reported largest effect sizes in the domains of attention, speed of information processing, working memory, verbal learning and memory (15, 33, 34, 39). Impaired performance on the Hinting Task in patients is in accordance with meta-analytic results on theory of mind in schizophrenia (40).

The recognition of happy affect is known to be relatively preserved in patients with psychosis (41), which was illustrated by our results. The absence of significant alterations in RST conflict reaction time should be interpreted in combination with alterations in RST conflict accuracy. Results suggest that patients have more problems modifying their behaviour in response to negative feedback because they do not adapt adequately to the reversal condition by taking relatively more time. This could be explained by diminished cognitive flexibility in schizophrenia, conceptualized as “the ability to coordinate attention and response to two or more ongoing tasks and to adaptively switch response strategies in accord with contextual demands” (42).

Results in siblings and parents indicate that alterations in the domains of verbal learning, processing speed, reasoning, and problem solving, working memory and acquired knowledge, are amongst the most promising cognitive intermediate phenotypes for schizophrenia, which is supported by the literature (5, 8). Effect sizes for these domains were mild to moderate and between-group comparisons with control subjects survived conservative Bonferroni correction. The average effect size in siblings was, however, relatively low (-0.18 SD, range -0.01 to -0.43 SD) compared to the literature. Meta-analyses in unaffected relatives have reported effect sizes of -0.37 (range -0.28 to -0.54) and -0.41 (SD=0.38) (5, 6), which is more in line with the average effect size of -0.34 (range +0.13 to -1.17) that was found in the present parent sample.

Although several studies have reported that siblings perform worse on different versions of the CPT, including the CPT-AX (5, 6, 43), the present study did not find significant differences in siblings and parents. It may be that the CPT-HQ did not sufficiently burden early aspects of stimulus encoding and perceptual analysis, resulting in a processing load that was too low to be sensitive in relatives (43, 44).

While the social cognitive tasks yielded significant performance alterations in parents, differences between siblings and control subjects for DFAR Total (P<0.02), DFAR Fearful
(P<0.03), DFAR Angry (P<0.02) and the Hinting Task (P<0.04) did not survive Bonferroni correction. Worse performance in parents compared to siblings was unexpected because the parent group has passed the main age period of risk for developing a psychotic disorder (45). It is possible that the sibling group may have been relatively healthy, as perhaps they share less risk genes with their affected relatives than the parent group. However, it should be taken into consideration that age differences between the parent group and the oldest control group may have inflated the effect sizes in parents, although post-hoc analyses in the control group did not show an effect of age on the Hinting Task (B=0.00005, P=0.99) and the DFAR angry faces (B=-0.09, P=0.23). Although few studies have actually focussed specifically on parents of patients, worse theory of mind performance in parents of schizophrenia patients compared to parents of control subjects has been reported before (46). Compromised social cognitive functioning in parents but not in siblings supports prior evidence that mentalizing impairment in schizophrenia may reflect general cognitive deficits or residual symptom expression rather than representing a specific trait marker (47, 48). Previous research has suggested that neurocognition and social cognition are distinct, yet correlated, domains in psychosis (49, 50). Social cognition may serve as a mediator between neurocognition and community functioning, acting sequentially on the same pathway (51, 52). It can therefore be speculated that siblings with neurocognitive alterations but no social cognitive alterations may display no reduction in community functioning, whereas parents – displaying both neurocognitive and social cognitive alterations – might show diminished community functioning. The validity of this hypothesis should be tested in future studies.

Although comparing mean performance on individual cognitive domains between patients, relatives and controls provides indispensable information about putative intermediate phenotypes (figure 1), it does not show how the cognitive changes are distributed over the subject groups. For example, a mild alteration on a cognitive subtest in siblings could be caused by a majority of siblings displaying mild alterations or by a severe alteration in only a small subgroup while the majority displays no alterations. For this reason, total impairment scores were calculated for each of the subject groups. With the cut-off ≤-1 SD from the control mean, fairly equal proportions of patients demonstrated no, mild, moderate or severe impairments, which corresponds to the concept of a continuum of neurocognitive functioning in patients with schizophrenia (29). The proportion of patients with a neurocognitive profile within the normal range falls within the 15 to 30% that was reported in a recent review (10). Moving the cut-off from ≤-1 SD to ≤-2 SD dramatically increases the rate of patients without cognitive impairment, emphasizing the relevance of using more than one criterion for cognitive impairment (53).

The proportions of mild, moderate, and severe cognitive impairment in siblings and parents are intermediate between patients and control subjects, which may represent a dose-response relationship for genetic load. With the ≤-1 SD cut-off, approximately 50% of the relatives display no alterations against 70% of the control subjects. In a study by Egan et al (53), the proportions of subjects without cognitive impairments (≤-1SD) were higher: 62-75% in siblings as compared with 77-91% in control subjects. Using ≤-2 SD as a cut-off, these rates rise substantially to around 90% in the relatives. This indicates that with a more conservative criterion of alteration, parents and siblings move away from their affected relatives to become almost indistinguishable from control subjects.
Some limitations should be taken into consideration when interpreting the results. Not all subjects had complete cognitive test scores. This was predominantly due to problems in computerized assessment and data storage, which has been previously reported in a multi-centre study (54). Patient status however also affected the number of missing test results. Therefore, it cannot be excluded that patients with more cognitive alterations were more likely to have missing data, resulting in attenuated effect sizes. Secondly, the issue that control subjects are often matched to the patient group, resulting in control groups of younger age than the parent or mixed-relative group, is a common concern in this field of research (6). In a meta-analysis on cognitive functioning in unaffected relatives, an overall Cohen’s $d=0.36$ was reported for studies with age-matched groups versus $d=0.48$ for those with non-age-matched groups (6). Covarying for age in the analyses, although appealing, is an inappropriate way of dealing with group differences, because age represents a defining group characteristic of parents versus patients and siblings (38). Therefore, age-and gender corrected $z$-scores were calculated in this study to account for age differences. Although this resulted in the best possible fit, optimal age-correction could not be achieved, as parents in the highest age group had a mean age of 54 against a mean age of 46 in the oldest control group. Higher mean age in parents would be most likely to inflate performance differences in speeded tasks such as Digit Symbol-coding and CPT Accuracy. Instead, the tasks that showed more alterations in parents compared to siblings were unspeeded tasks assessing social cognition, making a confounding effect of age less likely.

In conclusion, this study suggests that familial predisposition to psychotic disorder is associated with immediate verbal learning, processing speed, reasoning and problem solving, acquired knowledge and working memory, with modest effect sizes. Tasks assessing set-shifting ability and vigilance with low processing load did not differentiate relatives from controls. While half of the unaffected relatives may experience some degree of cognitive impairment, severe cognitive impairment seems to be restricted to a minority.

**REFERENCE LIST**

8. KÉRI S, JANKA Z. Critical evaluation of cognitive dysfunctions as endophenotypes
14. BRAND N, JOLLES J. Learning and retrieval rate of words presented auditorily and visually. J Gen Psychol. 1985;112:201-10.
33. DICKINSON D, RAMSEY ME, GOLD JM. Overlooking the obvious: A meta-analytic
comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. Arch Gen Psychiatry. 2007 May 1, 2007;64:532-42.


