

Neurocognitive deficits in schizophrenia. Are we making mountains out of molehills?

S. Moritz^{1*†}, J. P. Klein^{2†}, T. Desler¹, H. Lill¹, J. Gallinat¹ and B. C. Schneider¹

¹Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246, Hamburg, Germany

²Department of Psychiatry and Psychotherapy, Lübeck University, Ratzeburger Allee 160, 23538 Lübeck, Germany

Background. Most original studies and all meta-analyses conducted to date converge on the conclusion that patients with schizophrenia display rather generalized neurocognitive deficits. For the present study, we reopen this seemingly closed chapter and examine whether important influences, such as lack of motivation and negative attitudes towards cognitive assessment, result in poorer secondary neuropsychological performance.

Method. A sample of 50 patients with an established diagnosis of schizophrenia were tested for routine neurocognitive assessment and compared to 60 nonclinical volunteers. Before and after the assessment, subjective momentary influences were examined (e.g. motivation, concerns about assessment, fear about poor outcome) for their impact on performance using a new questionnaire called the *Momentary Influences, Attitudes and Motivation Impact (MIAMI) on Cognitive Performance Scale*.

Results. As expected, patients performed significantly worse than controls on all neurocognitive domains tested (large effect size, on average). However, patients also displayed more subjective momentary impairment, as well as more fears about the outcome and less motivation than controls. Mediation analyses indicated that these influences contributed to (secondary) poorer neurocognitive performance. Differences in neurocognitive scores shrank to a medium effect size, on average, when MIAMI scores were accounted for.

Conclusions. The data argue that performance on measures of neurocognition in schizophrenia are to a considerable extent due to secondary factors. Poor motivation, fears and momentary impairments distinguished patients from controls and these variables heavily impacted performance. Before concluding that neurocognitive deficits in psychiatric patients are present, clinicians should take these confounding influences into account. Although patients with schizophrenia achieved, on average, worse test scores than controls, a large subgroup displayed spared performance.

Received 29 November 2016; Revised 20 February 2017; Accepted 21 March 2017; First published online 9 May 2017

Key words: Effort, motivation, neurocognition, psychosis, schizophrenia.

Introduction

Neurocognitive deficits, such as impairments of memory and attention, are neither core criteria for the diagnosis of schizophrenia (APA, 2013) nor pathognomonic (Moritz *et al.* 2001, 2002; Bora *et al.* 2009; Palmer *et al.* 2009). Yet, this research domain has attracted much interest since the times of Kraepelin (Kraepelin, 1899), who claimed that the disorder (then called *dementia praecox*) is characterized by neurocognitive decline ultimately manifesting as dementia. The vast majority of original studies, narrative reviews and meta-analyses arrive at the conclusion that patients with schizophrenia

have major neurocognitive deficits (Heinrichs & Zakzanis, 1998; Keefe *et al.* 1999; Reichenberg, 2010; Schaefer *et al.* 2013; Fatouros-Bergman *et al.* 2014). As a rule of thumb, deficits among individuals with schizophrenia are estimated to lie approximately 1 standard deviation below the norm. However, unlike Kraepelin's assumption, these deficits are rather stable and there is no evidence for progressive neurodegenerative decline subsequent to the first episode (Mesholam-Gately *et al.* 2009; Bora & Murray, 2014).

Neurocognitive deficits, which have been often detected in comparison to samples of demographically matched controls, are mainly attributed to the primary disorder (i.e. schizophrenia). They are hypothesized to act as a risk factor for psychosis, particularly the negative and/or disorganized syndrome, and a more severe course of illness (Brewer *et al.* 2006; Carlsson *et al.* 2006). Since the mid-1990s, a number of studies (Green, 1996; Green *et al.* 2000, 2004) have suggested that objective neurocognitive deficits, in conjunction

* Address for correspondence: S. Moritz, Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246, Hamburg, Germany.

(Email: moritz@uke.de)

† SM and JPK have equally contributed to the manuscript and share first authorship.

with (and sometimes overshadowed by) impairments in social cognition (Fett *et al.* 2011), predict poor functional outcome and may also compromise treatment outcome for some types of intervention (Kurtz, 2011). Consequently, efforts to develop trainings (Wykes *et al.* 2011; Cella *et al.* 2015) and medications (López-Muñoz & Álamo, 2011) that augment cognitive functioning (i.e. cognitive enhancers) are on the rise.

In this article, we will argue that the literature has neglected important influences that may contribute to neurocognitive test deficits, and thus, prohibit a fair comparison between patients and controls. These factors have only recently been addressed systematically (Fervaha *et al.* 2014; Strauss *et al.* 2015). For example, most patients with schizophrenia are prescribed medication, particularly antipsychotic agents, which – with the possible exception of clozapine – may induce Parkinsonian symptoms. Such symptoms will compromise performance in any motor test, such as the Trail Making Test (Fervaha *et al.* 2015), whether or not cognitive faculties are affected. Additionally, patients are also often given tranquilizers and anticholinergic drugs known to decrease memory and attention (Vinogradov *et al.* 2009). Although it is implicitly claimed that neurocognitive deficits (as manifestation of brain-related dysfunction) underlie negative symptoms (Hovington & Lepage, 2012), such as avolition and poor motivation, the causality may in fact be the reverse.

A number of studies suggest that patients with schizophrenia are less willing to engage to the best of their abilities (Fervaha *et al.* 2014), which makes it almost impossible to gain a true picture of their neurocognitive potential. Unlike psychophysiological or laboratory measures (e.g. blood tests), scores on neurocognitive tests heavily rely on the cooperation of the patient. In addition, patients may be distracted by rumination, unusual ideas and/or hallucinations during testing, which could direct attention away from the task at hand. Matching samples for age, gender and education is thus not fully sufficient. Finally, a restricted non-arousing hospital environment itself may lead to a decrement in performance, as has been shown in non-psychiatric populations (Wilson *et al.* 2012; Mathews *et al.* 2014). In patients with psychosis, the relationship between hospitalization and cognition is complex and perhaps mediated by more severe symptomatology (Harvey *et al.* 2013).

We have previously investigated secondary influences on neuropsychological functioning in patients with obsessive-compulsive disorder (OCD) (Moritz *et al.* 2012), for whom neurocognitive deficits have also been described (Abramovitch *et al.* 2013; Shin *et al.* 2014) – albeit to a lesser degree than in individuals with schizophrenia. Poor motivation and

checking compulsions were significantly associated with cognitive performance deficits. We concluded that while the study does not rule out that neurocognitive deficits may play a role in the formation of OCD symptoms, the reverse relationship should be considered as well. Likewise, another study found that in depression, cognitive deficits are also aggravated by motivational deficits (Scheurich *et al.* 2008).

In the present study, we examined whether the rationale of the aforementioned studies on OCD and depression can also be applied to patients with schizophrenia and compared a sizable sample of patients with schizophrenia to healthy controls. We assessed all individuals both before and after comprehensive neurocognitive testing with a newly developed self-report scale aimed at uncovering several subjective influences on cognitive performance, including effort, motivation and other factors (i.e. concerns about assessment, fear about poor outcome, poor performance due to momentary influences). Assessing patients before the assessment was deemed important as post-assessment data may be biased by tactical responses (e.g. a self-serving bias to blame poor performance on low motivation). We expected that patients would perform worse on neurocognitive tests than controls as a group but that differences would be (partially) mediated by other group differences, particularly on motivation and fears about the outcome of testing. While we expected that the nonclinical controls would be highly motivated to produce good results, we hypothesized that patients with schizophrenia would show less ambition.

Methods

The clinical sample was comprised of 50 inpatients consecutively referred to the clinical neuropsychology unit of the Department of Psychiatry and Psychotherapy at the University Medical Center in Hamburg-Eppendorf (Germany). All patients met diagnostic criteria for a schizophrenia spectrum disorder as verified by a psychiatrist. All patients were prescribed psychotropic medication, mainly atypical antipsychotics. The total score of the Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1962) was 41.29 (18.25), which corresponds to ‘moderately ill’ according to the Clinical Global Impression (CGI) scale (Leucht *et al.* 2005). On average, patients had 1.51 (1.13) prior inpatient admissions (range: 0–7). A group of 60 people served as nonclinical controls, for whom a psychiatric diagnosis had been excluded by means of a short psychiatric screening for mental disorders based on the MINI International Neuropsychiatric Interview, which had been adapted to *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* (APA, 2013) criteria by the authors (MINI)

(Sheehan et al. 1998). Control participants were sought from various sources via word-of-mouth and advertisements. General exclusion criteria were major neurological disorders and age younger than 18 or older than 65 years.

All participants underwent a battery of neuropsychological tests, which took approximately 2–3 h depending on participants' speed and requests for breaks. Prior to the assessment and afterwards, participants were asked to complete a short questionnaire asking for subjective momentary attitudes, momentary impairments and motivation (see below).

Characteristics of patients and controls are presented in Table 1.

Participants completed the following battery of neurocognitive tasks:

Trail Making Test (TMT) (Reitan, 1992)

Psychomotor speed was assessed with the TMT Part A (adult version). The individual was asked to connect numbers in ascending order as quickly as possible. Part B assesses set-shifting and requires alternating between numbers and letters, again in ascending order. Age-adapted standard scores were applied (Tombaugh, 2004).

Wisconsin Card Sorting Test (WCST) (Heaton, 1981)

Participants completed a computerized version of the WCST (Loong, 1990), which followed the procedure of the original paper test. Participants were shown a maximum of 128 cards, which had to be matched according to three varying principles (i.e. number of items, color, shape). These matching principles were not disclosed to the individual. Via a tone (high *v.* low) and a corresponding verbal cue, it was indicated whether a match was correct or not. Categories completed and perseverative errors served as dependent variables.

Auditory Verbal Learning Test (AVLT) (Lezak, 1995)

Memory and learning were assessed with the German version of the AVLT (Helmstaedter et al. 2001). A list of 15 words (List A) was read to the participants five times. After each trial, the individual had to repeat as many words as possible in no particular order. Afterwards, the individual had to memorize items from a separate inference list (List B). Following this, words from List A had to be recalled without renewed presentation. Thirty minutes later, the individual was asked to repeat words from List A only. The AVLT allows for the computation of several memory parameters (Lezak, 1995). Learning was indexed by the sum of correctly recalled words on Trials 1 to 5. The number of correctly reproduced words after the 30-min delay served as an

index for long-term memory. Normative scores for the German version of the task are available for different age ranges (Helmstaedter et al. 2001).

Selective attention Test d2 (Brickenkamp, 1962)

Test d2 is a letter cancellation test that taps selective attention. The individual had to cross out the letter *d* whenever it was presented with two small lines; *d*'s with more or less than two lines, or any stimuli containing the character *p*, served as distracters. Following a practice trial, 14 rows containing target and distracter stimuli were presented. Participants had to complete each row within 20 s. The test is scored for errors and number of correctly crossed out stimuli within the allotted time. Normative scores for the parameter *concentration performance* (*Konzentrationsleistung*) served as the primary parameter. Age-adjusted normative scores were derived from a large population sample (Schmidt-Atzert, 2004).

Divided attention subtest from the Test for Attentional Performance (TAP) (Zimmermann & Fimm, 1995)

For divided attention, participants had to perform two different tasks concurrently on a computer. The space bar had to be pressed whenever asterisks formed a rectangle on a 4 × 4 dot matrix (optical target), or whenever two tones of the same frequency were repeated (acoustic target). One hundred optical and 200 acoustic trials were presented. There were 16 targets for each modality. Norm values derived from a large sample of participants are available for median reaction times and number of omissions (Zimmermann & Fimm, 1995).

Selective attention subtest (Go-NoGo) from the Test for Attentional Performance (TAP) (Zimmermann & Fimm, 1995)

Participants were presented with either an 'x' (target, 'go') or a '+' (distracter, 'nogo') in this simple computerized task of selective attention. A button had to be pressed whenever a target stimulus appeared. Norm scores are available for speed variables (Zimmermann & Fimm, 1995).

Story Recall from the Wechsler Memory Scale (WMS) – Revised Edition (Woodard & Axelrod, 1987)

Two short stories were read and the participant was asked to repeat as much of the story as they could remember immediately after the completion of each story. Thirty minutes later, participants were again asked to recall as much of the story as possible. No other tests of verbal memory were presented during

Table 1. Differences between schizophrenia and healthy individuals on demographic characteristics and scores on the MIAMI subscales

Variables	Schizophrenia (n = 50)		Healthy (n = 60)		Statistics			
	n	%	n	%	χ^2	df	p	
<i>Demographic characteristics</i>								
Sex (female/male)	33/17	66%/34%	31/29	52%/48%	2.30	1	0.131	–
	M	S.D.	M	S.D.	$ t /\chi^2$	df	p	d
Age in years	30.84	9.78	32.35	10.81	0.76	107	0.450	0.15
Years of formal education	11.77	1.70	12.03	1.44	0.87	106	0.387	0.17
<i>MIAMI pre</i>								
Low motivation	2.12	0.47	1.99	0.36	1.60	108	0.113	0.30
Concerns about assessment	2.23	0.59	1.98	0.41	2.54	108	0.013	0.49
Fear about poor outcome	2.34	0.64	1.75	0.55	5.15	108	<0.001	0.98
Unfavorable momentary influences	2.36	0.55	1.71	0.45	6.88	108	<0.001	1.30
Total score	2.26	0.35	1.86	0.27	6.87	108	<0.001	1.30
<i>MIAMI post</i>								
Low motivation	1.74	0.54	1.40	0.47	3.53	108	<0.001	0.67
Concerns about assessment	1.79	0.43	1.54	0.38	3.28	108	0.001	0.62
Fear about poor outcome	2.18	0.56	1.75	0.44	4.58	108	<0.001	0.87
Unfavorable momentary influences	2.11	0.48	1.58	0.42	6.25	108	<0.001	1.19
Total score	1.96	0.33	1.57	0.30	6.50	108	<0.001	1.24

Note. MIAMI = Momentary Influences, Attitudes and Motivation Impact (MIAMI) on Cognitive Performance Scale, higher scores designate more negative thoughts and feelings associated with the assessment.

the 30-min follow-up time. Scoring followed instructions in the manual; German norm values were applied (Härting *et al.* 2000).

Similarities from the Wechsler Adult Intelligence Scale, 4th edition (Jacobs & Petermann, 2007; Wechsler, 2008)

In this verbal reasoning task, the individual has to deduce what two words have in common. Scoring followed the original manual. Scaled scores exist for a large German population (Jacobs & Petermann, 2007).

Matrix Reasoning from the Wechsler Adult Intelligence Scale, 4th edition (Jacobs & Petermann, 2007; Wechsler, 2008)

The matrix subtest measures nonverbal reasoning. The individual is presented with a pattern sequence and must select the item that completes the sequence from five response choices. Scaled scores exist for a large German population.

Zoo Test from the Behavioural Assessment of the Dysexecutive Syndrome (BADS) (Wilson *et al.* 2003)

The zoo test captures planning ability. The individual had to plan a route to visit six of 12 possible locations in a zoo. The test has two phases (i.e. first, a demanding situation with little external structure was given;

then, a concrete, externally imposed strategy had to be followed).

Block Design from the Wechsler Adult Intelligence Scale 4th edition (Jacobs & Petermann, 2007; Wechsler, 2008)

In this visuospatial performance test, participants must re-create two-dimensional patterns using colored blocks. Task demands become increasingly difficult over time. Scoring is made according to accuracy and time following instructions in the manual. Scaled scores exist for a large German population.

In keeping with prior studies, we expected no clear profile of impairment across domains. Therefore, we composed two new cognitive subscores, which aggregated either speed or performance variables, expressed as z-values, rather than proceeding with individual domains (if higher scores designated better performance, as in the memory tests, parameters were reversed).

Assessment of subjective influences on cognitive performance

Prior and subsequent to the cognitive assessment, participants received either the baseline or post-assessment version of a questionnaire called the *Momentary Influences, Attitudes and Motivation Impact (MIAMI) on Cognitive Performance Scale*. Each version

consists of 18 items. The baseline version covers four domains[†]: (1) poor motivation [e.g. 'I am not motivated at all to take part in the assessment.'; 'I am willing to do my best.' (reverse)], (2) concerns about the assessment [e.g. 'I experience the situation as very unpleasant and would like to leave.'; 'I know what I can expect from the assessment.' (reverse)] (3) fear about poor outcome [e.g. 'I worry that the tasks will be too difficult for me.'; 'In general, I feel confident before and after assessments. I am aware of my cognitive strengths and am able to show them.' (reverse)] (4) negative momentary influences [e.g. 'Right now, I feel very tired and exhausted.'; 'I feel fit and capable.' (reverse)]. Items were rated on a four-point Likert scale (1 – fully agree, 2 – rather agree, 3 – rather disagree, 4 – disagree). Scores were inverted for all subscales; higher scores designate more negative thoughts and feelings associated with the assessment.

Following the neuropsychological assessment, participants completed the post version of the MIAMI questionnaire, which again consisted of 18 items that were rated using the same Likert scale as before. The retrospective ratings measure the same domains albeit with different items: (1) poor motivation [e.g. 'I was not motivated and accordingly did not achieve my best performance.'; 'I pushed myself and gave my best.' (reverse)], (2) concerns about the assessment [e.g. 'I felt under pressure by the test situation.'; 'The testing situation was not as bad as I had expected.' (reverse)] (3) fear about poor outcome (e.g. 'I am fearful about the results.') (4) negative momentary influences [e.g. 'I had a headache during the assessment.'; 'I could sufficiently concentrate during the assessment' (reverse)]. Subscale scores were inverted; higher scores designate more negative thoughts and feelings associated with the assessment. In an unrelated study (Moritz et al. submitted) in 30 patients with OCD, the MIAMI total score modestly correlated ($r=0.42$, $p<0.01$) with an expert rating (the assessor was asked to determine the patient's motivation and distraction by symptoms among other indexes).

Strategy of data analysis

We first compared the two groups using *t*-tests for independent samples on all neurocognitive tasks. In order to assess whether momentary influences, assessment-related fears and concerns, as well as motivation, as tapped by the MIAMI scale, mediate group differences on neuropsychological functioning, we adopted a two-fold strategy. For the main analyses, we

calculated mediation analyses using Hayes' *process* procedure (model 4; 5000 bootstrap samples) with group as the independent variable (*x*), neuropsychological functioning as the dependent variable (*y*) and MIAMI total scores as the mediator (*M*). Analyses were run separately for the two neuropsychological indexes (speed, performance) and two MIAMI (pre, post) total scores. Second, we recalculated mean comparisons by entering MIAMI total scores (pre and post separately) as covariates to explore whether effect sizes would be attenuated substantially. Please note that the latter analyses are meant to aid in the interpretation of the mediation analyses and represent no proper test of whether the mediator/covariate exerts a significant effect.

Results

Both samples were very similar with respect to age, gender and education. Patients differed from controls on all MIAMI subscores and total scores, except for the motivation subscale at baseline. Cronbach's alphas for the MIAMI total score were 0.70 (baseline) and 0.78 (post). As expected, patients performed worse on most cognitive test parameters (16 out of 18 scores) relative to controls (see Table 2). On average, between group-effect sizes were large ($M_d=0.78$). Almost two out of three patients (64%) performed at least one standard deviation below the mean on two or more tests (controls: 32%)². When stricter criteria were applied (at least two standard deviations below the mean on two or more tests, for a critical discussion of thresholds see Abramovitch & Schweiger, 2015) the percentage dropped to 14% in patients *v.* 0% in controls.

As shown in Table 3, correlational analyses indicated significant associations between aggregated neuropsychological and MIAMI scores in 17 of 20 correlations (2 aggregated neuropsychological scores \times 2 (pre, post) \times 5 MIAMI (sub)scales); the average absolute correlation was $r=0.30$. Only motivation prior to testing, as well as concerns about the assessment (post), were uncorrelated with neurocognitive functioning.

Figure 1 shows that, with the exception of mediation B (i.e. the effect of group on speed, mediator: MIAMI post), the significant total relationship between group and neurocognition was largely mediated by MIAMI total scores, as evidenced by a significant Sobel test (in addition, confidence intervals of the indirect effect did not cross zero), even though all direct effects remained significant. Results remained unchanged when the subscale *Negative Momentary Influences* was removed from the total score, as one may argue that this subscale may partly reflect the subjective perception of one's actual cognitive impairment.

[†] The notes appear after the main text.

Table 2. Differences in neurocognitive functioning between schizophrenia and healthy individuals

Variables/somains	Schizophrenia (n = 50)			Healthy (n = 60)			% 1s.d. ≤ Norm	Statistics				ANCOVA (MIAMI pre)		ANCOVA (MIAMI post)	
	M	s.d.	Norm	M	s.d.	Norm		t /χ ²	df	p	d	p	d	p	d
<i>Speed</i>															
Trail-Making Test A in sec.	29.77	10.42	39.41 ^P	22.02	7.90	64.70 ^P	20%/8%	4.44	108	<0.001	0.84	0.009	0.51	0.004	0.57
Divided attention auditory in ms	651.14	100.39	38.90 ^T	602.35	81.32	42.65 ^T	60%/38%	2.80	107	0.006	0.53	0.248	0.23	0.031	0.42
Divided attention visual in ms	807.98	107.70	46.36 ^T	764.17	100.40	50.52 ^T	20%/12%	2.20	108	0.030	0.42	0.300	0.20	0.078	0.34
Go-nogo in ms	431.08	96.55	44.88 ^T	418.05	83.35	48.27 ^T	38%/30%	0.76	108	0.449	0.14	0.665	0.03	0.977	0.00
<i>Attention</i>															
Go-nogo errors	1.02	2.14	–	0.83	1.30	–	–	0.57	107	0.572	0.11	0.661	0.03	0.929	0.00
d2 concentration index	135.22	35.22	27.28 ^P	189.02	46.88	71.05 ^P	36%/22%	6.69	108	<0.001	1.30	<0.001	0.87	<0.001	0.92
divided attention omissions	3.44	3.57	40.36 ^T	1.18	1.58	48.72 ^T	56%/20%	4.15	64.84	<0.001	0.82	0.012	0.49	0.013	0.49
<i>Memory</i>															
AVLT learning	50.20	13.30	47.26 ^T	61.32	7.76	58.60 ^T	34%/2%	5.22	75.66	<0.001	1.02	<0.001	0.72	0.001	0.64
AVLT retention	10.30	3.49	45.46 ^T	13.28	2.21	56.65 ^T	38%/5%	5.23	79.83	<0.001	1.02	<0.001	0.72	<0.001	0.72
Logical Memory immediate	24.28	10.31	32.85 ^P	32.15	5.79	68.27 ^P	40%/11%	4.69	68.31	<0.001	0.94	0.001	0.64	0.001	0.64
Logical Memory delayed	19.17	8.71	31.65 ^P	28.95	6.07	68.58 ^P	44%/11%	6.55	78.74	<0.001	1.30	<0.001	0.94	<0.001	0.97
<i>Reasoning</i>															
Similarities	24.55	5.40	8.92 ^s	27.55	3.47	10.70 ^s	31%/7%	3.36	78.46	0.001	0.66	0.012	0.49	0.002	0.64
Matrix reasoning	19.16	4.70	10.12 ^s	22.28	2.60	12.33 ^s	16%/0%	4.19	73.29	<0.001	0.82	0.004	0.57	0.008	0.52
<i>Spatial Performance</i>															
Block design	46.92	12.70	9.48 ^s	56.88	8.22	12.47 ^s	30%/0%	4.78	81.00	<0.001	0.93	0.001	0.64	<0.001	0.70
<i>Executive Functioning</i>															
Trail-Making Test B in sec.	74.67	35.77	32.45 ^P	51.15	14.60	59.33 ^P	39%/12%	4.32	61.03	<0.001	0.86	0.004	0.58	0.002	0.62
WCST categories	4.54	2.04	–	5.57	1.11	–	–	3.13	68.80	0.003	0.62	0.064	0.36	0.123	0.30
WCST perseveration	18.47	13.70	–	10.46	5.49	–	–	3.51	47.43	0.001	0.77	0.007	0.56	0.046	0.41
Zoo Test Total	13.71	3.39	–	15.83	0.83	–	–	4.27	52.67	<0.001	0.86	0.006	0.55	0.002	0.64
<i>Aggregated Scores (z-transformed)</i>															
Speed (Slowing)	0.30	0.72	–	–0.24	0.64	–	–	4.17	108	<0.001	0.79	0.04	0.40	0.006	0.54
Accuracy (Malperformance)	0.40	0.75	–	–0.32	0.35	–	–	6.32	64	<0.001	1.23	<0.001	0.84	<0.001	0.86

Notes. T = T-scores (M = 50), P = percentile (M = 50), S = scaled score (M = 10).

Table 3. Correlations between MIAMI subscales with the aggregated neuropsychological scores (full sample)

MIAMI scores	Low speed	Poor accuracy
<i>Baseline</i>		
Low motivation	0.062	-0.016
Concerns about assessment	0.348 (****)	0.353 (****)
Fear about poor outcome	0.260 (**)	0.382 (****)
Unfavorable momentary influences	0.364 (****)	0.435 (****)
Total score	0.403 (****)	0.466 (****)
<i>Post</i>		
Low motivation	0.251 (**)	0.232 (*)
Concerns about assessment	0.204 (*)	0.160 (+)
Fear about poor outcome	0.227 (*)	0.489 (****)
Unfavorable momentary influences	0.198 (*)	0.451 (****)
Total score	0.301 (****)	0.466 (****)

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, **** $p < 0.001$.

The result of the mediation analysis was corroborated by the ANCOVA analyses displayed in the last columns of Table 2, showing that effect sizes decreased substantially if MIAMI scores were accounted for (for ANCOVAs using the MIAMI baseline score, the average effect size dropped to $d = 0.53$, for the post version it was $d = 0.51$; uncorrected average effect size: $d = 0.78$, see above). The uncorrected group comparisons yielded twelve large ($d \geq 0.8$) and 5 medium effects ($d \geq 0.5$; 18 subscores and 2 aggregated scores). When the MIAMI baseline total score was accounted for, only 3 comparisons showed large effects and 9 effects were in the medium range. For the MIAMI post score (i.e. retrospective assessment of the test situation), a similar picture emerged (i.e. 3 large effects, 10 medium effects).

Discussion

At first glance, our results corroborate a plethora of studies suggesting that patient with schizophrenia have severe neurocognitive deficits (Heinrichs & Zakzanis, 1998; Keefe et al. 1999; Schaefer et al. 2013; Fatouros-Bergman et al. 2014). On average, we found large effect sizes indicating better neurocognitive performance among controls. However, upon closer inspection, some caution is warranted when interpreting these findings.

First, only a minority of patients in our study had neurocognitive abnormalities according to standard criteria (Abramovitch & Schweiger, 2015); 14% of patients scored at least two standard deviations below the mean on two or more tests. This rose to

64% if the threshold is set very liberally (at least one standard deviation below the norm on two or more tests). This result is clearly not novel or surprising. Clinicians trained in statistics know that group differences eclipse outliers to both ends of the performance distribution. Yet, when such findings are communicated to peers and the public, researchers often fail to highlight that results cannot be generalized to every single individual, thereby possibly increasing stigma and fears (e.g. about dementia). Of note, even Kraepelin (1899) described a substantial subgroup of patients who recovered without residual symptoms, casting doubt on his general claim about premature dementia, which continues to fuel the interest in cognitive impairment in schizophrenia.

Secondly, patients markedly differed from controls regarding their endorsements of subjective influences on cognitive performance before and after assessment, particularly regarding fears about the outcome of the test, distracting symptoms (e.g. headache, tiredness), and performance motivation³ (see also Fervaha et al. 2014). Indeed, as predicted, a number of these variables heavily impacted performance. If taken into account statistically, group differences in neurocognitive functioning were still detectable (= molehill) but not as huge (= mountain) as before – after correcting for subjective influences on cognitive performance, differences declined from large effects to medium. Differences in subjective influences on cognitive performance found before testing are especially meaningful, as these suggest that differences were real and not just ad-hoc explanations for poor performance. Again, while neurocognitive deficits were detectable among patients with schizophrenia, even if contextual factors are accounted for, this says little about individual performance. Momentary influences, negative attitudes and fear about the assessment may shift performance of an unimpaired participant to *impaired* or push poor performances in the range of one standard deviation below the norm to two standard deviations (which is often considered diagnostically relevant and may thus lead to additional diagnoses).

We argue that the investigation has a number of important implications. As researchers, we might need to adjust our language. While it is circumstantial to talk about ‘a subgroup of patients’ rather than patients (in general), the latter formulation enhances stigma as not all patients have neurocognitive deficits. In light of the high prevalence of depression, suicidal ideation and low self-esteem in those afflicted with schizophrenia, we must be careful not to insult patients when confronting them with test results or to induce further fears that may propagate self-stigma (e.g. giving a comorbid diagnosis of mild cognitive impairment), particularly when transient secondary

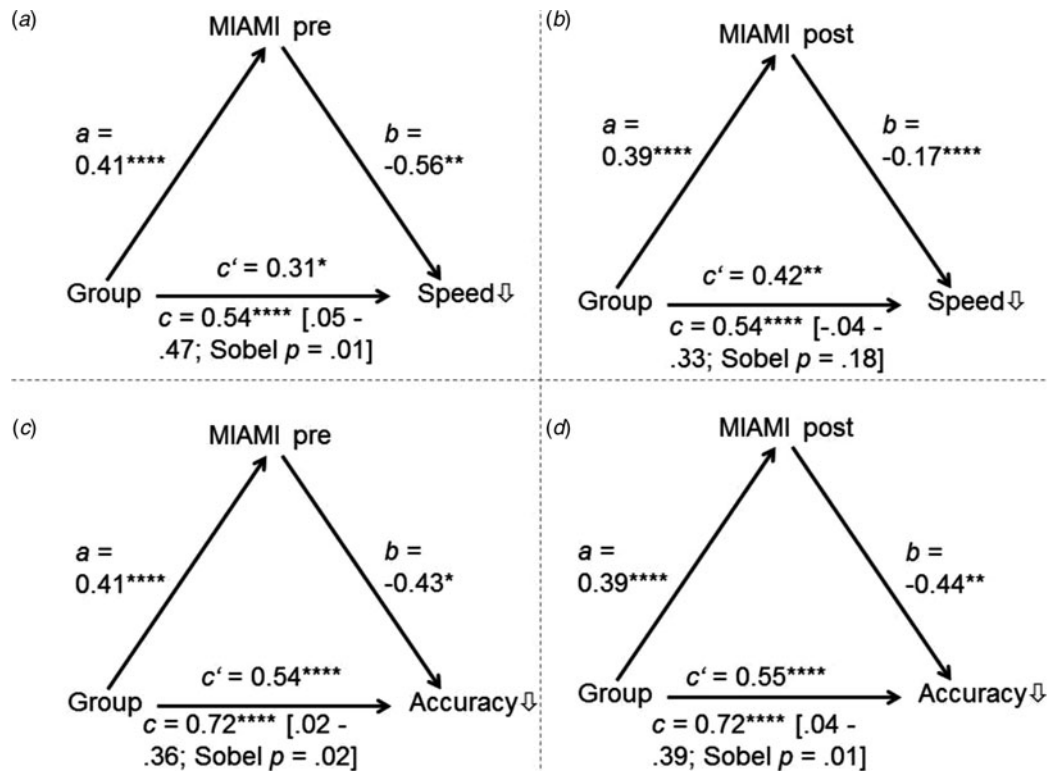


Fig. 1. Results from the mediation analyses using Hayes' process macro. Group status predicted MIAMI scores ($=a$) and these in turn predicted neuropsychological measures ($=b$). Except for analysis B, the indirect effect was significant (i.e. confidence intervals did not cross 0) indicating that momentary influences, concerns and fears as well as motivation (MIAMI subscales) largely contribute to group differences (0 = healthy, 1 = patients) on test results.

influences, like medication and poor motivation cannot be ruled out. As shown, due to internal (e.g. poor motivation due to a self-fulfilling prophecy), external (e.g. medication) or symptom-related (e.g. rumination, voice-hearing that distracts the individual) factors, many patients do not achieve their best performance. We need to pay closer attention to these factors before making serious recommendations; for example, whether to further academic education or career decisions. As mentioned before, matching control samples for demographic variables is clearly not enough in this regard.

We would also like to highlight a number of limitations. First, results need to be replicated with larger samples and psychiatric control groups, for whom similar findings have been reported (Scheurich *et al.* 2008; Moritz *et al.* 2012). One may also object that our samples were equated for background characteristics, such as education, which is usually lower in patients than controls. We also excluded patients with dependence disorders or neurological problems, which are also frequently observed in patients with schizophrenia. We did this so that we could attribute any differences on cognitive test results to diagnostic

status and not to secondary factors other than MIAMI scores. This strategy may have contributed to a slight overestimation of patients' performance. Our results thus need replication in more severe samples. Second, motivation, attitudes and momentary symptoms were measured with self-report scales. Subjective poor performance may prompt a self-serving bias and a tendency to blame performance on context, motivation or other factors. Indeed, a tendency to externalize blame has been frequently observed in patients with psychosis (Bentall *et al.* 1994; Kinderman & Bentall, 1996). In order to minimize such biases, we administered the MIAMI questionnaire both before and after assessment. Subsequent trials may also try to collect objective data (e.g. number of yawns during assessment, psychophysiological recording) and explore the impact of specific symptoms (e.g. distraction due to voice-hearing, delusional pre-occupations etc.). In addition, with the exception of one study (see methods section), the MIAMI scale has not been validated against other measures of motivation or performance anxiety. We could also not control for medication in our analysis as all patients were medicated; chlorpromazine equivalent

or percentage of maximum dosage algorithms cannot fully address this problem (Strauss *et al.* 2015).

For the future, it seems important to examine if secondary impairments may also depend on how investigations are framed. For example, some investigators explicitly tell patients that tests are carried out to investigate impaired cognition, which is introduced as an important factor for the formation of the illness. Such information may elicit *stereotype threat* (Pennington *et al.* 2016), frustrate patients and impede their full engagement (Kit *et al.* 2008). Unfortunately, studies often report little information beyond which specific tests were administered. From studies with women and older people (Spencer *et al.* 1999), we already know that even subtle formulations (framing) that potentially trigger stigma can impair performance.

To conclude, bold claims that 'neurocognitive impairment is present in most, if not all, persons with schizophrenia' (Palmer *et al.* 2009) (p. 365) may need to be revised in light of our findings. We should refrain from monocausal inferences when low test scores are detected in patients. While brain-related deficits may undoubtedly play a vital role, they are not the only source of impairment.

Notes

- ¹ For a copy of the MIAMI questionnaires, please write to moritz@uke.de
- ² We would like to thank one anonymous referee for suggesting this algorithm.
- ³ Interestingly, differences in motivation emerged for the post but not the baseline assessment. Based on clinical experience, we speculate that many patients are initially as motivated as controls but lose interest and motivation in the course of the assessment.

References

- Abramovitch A, Abramowitz JS, Mittelman A** (2013). The neuropsychology of adult obsessive-compulsive disorder: a meta-analysis. *Clinical Psychology Review* **33**, 1163–1171.
- Abramovitch A, Schweiger A** (2015). Misuse of cognitive neuropsychology in psychiatry research: the intoxicating appeal of neo-reductionism. *Behavior Therapist* **38**, 187–191.
- American Psychiatric Association** (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, 5th edn. American Psychiatric Association: Arlington, VA, USA.
- Bentall RP, Kinderman P, Kaney S** (1994). The self, attributional processes and abnormal beliefs: towards a model of persecutory delusions. *Behaviour Research and Therapy* **32**, 331–341.
- Bora E, Murray RM** (2014). Meta-analysis of cognitive deficits in ultra-high risk to psychosis and first-episode psychosis: do the cognitive deficits progress over, or after, the onset of psychosis? *Schizophrenia Bulletin* **40**, 744–755.
- Bora E, Yucel M, Pantelis C** (2009). Cognitive functioning in schizophrenia, schizoaffective disorder and affective psychoses: meta-analytic study. *British Journal of Psychiatry* **195**, 475–482.
- Brewer WJ, Wood SJ, Phillips LJ, Francey SM, Pantelis C, Yung AR, Cornblatt B, McGorry PD** (2006). Generalized and specific cognitive performance in clinical high-risk cohorts: a review highlighting potential vulnerability markers for psychosis. *Schizophrenia Bulletin* **32**, 538–555.
- Brickenkamp R** (1962). *Aufmerksamkeits-Belastungs-Test Handanweisung d-2 [Attention-Pressure-Test Manual d2]*. Hogrefe: Göttingen.
- Carlsson R, Nyman H, Ganse G, Cullberg J** (2006). Neuropsychological functions predict 1- and 3-year outcome in first-episode psychosis. *Acta Psychiatrica Scandinavica* **113**, 102–111.
- Cella M, Reeder C, Wykes T** (2015). Cognitive remediation in schizophrenia-now it is really getting personal. *Current Opinion in Behavioral Sciences* **4**, 147–151.
- Fatouros-Bergman H, Cervenka S, Flyckt L, Edman G, Farde L** (2014). Meta-analysis of cognitive performance in drug-naïve patients with schizophrenia. *Schizophrenia Research* **158**, 156–162.
- Fervaha G, Agid O, Takeuchi H, Lee J, Foussias G, Zakzanis KK, Graff-Guerrero A, Remington G** (2015). Extrapyramidal symptoms and cognitive test performance in patients with schizophrenia. *Schizophrenia Research* **161**, 351–356.
- Fervaha G, Zakzanis KK, Foussias G, Graff-Guerrero A, Agid O, & Remington G** (2014). Motivational deficits and cognitive test performance in schizophrenia. *JAMA Psychiatry* **71**, 1058–1065.
- Fett AKJ, Viechtbauer W, Dominguez MDG, Penn DL, van Os J, Krabbendam L** (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neuroscience and Biobehavioral Reviews* **35**, 573–588.
- Green MF** (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry* **153**, 321–330.
- Green MF, Kern RS, Braff DL, Mintz J** (2000). Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the 'right stuff'? *Schizophrenia Bulletin* **26**, 119–136.
- Green MF, Kern RS, Heaton RK** (2004). Longitudinal studies of cognition and functional outcome in schizophrenia: implications for MATRICS. *Schizophrenia Research* **72**, 41–51.
- Härting C, Markowitsch HJ, Neufeld H, Calabrese P, Deisinger K, Kessler J** (2000). *Wechsler Gedächtnistest - Revidierte Fassung (WMS-R) [Wechsler Memory Scale - revised version (WMS-R)]*. Hans Huber: Bern.
- Harvey PD, Loewenstein DA, Czaja SJ** (2013). Hospitalization and psychosis: Influences on the course of cognition and everyday functioning in people with schizophrenia. *Neurobiology of Disease* **53**, 18–25.
- Heaton RK** (1981). *Wisconsin Card Sorting Test. Manual*. Psychological Assessment Resources: Odessa, FL.

- Heinrichs RW, Zakzanis KK (1998). Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology* **12**, 426–445.
- Helmstaedter C, Lendt M, Lux S (2001). *VLMT Verbaler Lern- und Merkfähigkeitstest [VLMT Verbal Learning and Memory Test]*. Belz Test: Göttingen.
- Hovington CL, Lepage M (2012). Neurocognition and neuroimaging of persistent negative symptoms of schizophrenia. *Expert Review of Neurotherapeutics* **12**, 53–69.
- Jacobs C, Petermann F (2007). Wechsler Intelligenztest für Erwachsene (WIE) [Wechsler Intelligence Scale for Adults]. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie* **55**, 205–208.
- Keefe RS, Silva SG, Perkins DO, Lieberman JA (1999). The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: a review and meta-analysis. *Schizophrenia Bulletin* **25**, 201–222.
- Kinderman P, Bentall RP (1996). A new measure of causal locus: the internal, personal and situational attributions questionnaire. *Personality and Individual Differences* **20**, 261–264.
- Kit KA, Tuokko HA, Mateer CA (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review* **18**, 132–148.
- Kraepelin E (1899). *Psychiatrie. Ein Lehrbuch für Studierende und Aerzte [Psychiatry. A textbook for students and physicians]*. J. A. Barth: Leipzig.
- Kurtz MM (2011). Neurocognition as a predictor of response to evidence-based psychosocial interventions in schizophrenia: what is the state of the evidence? *Clinical Psychology Review* **31**, 663–672.
- Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R (2005). Clinical implications of Brief Psychiatric Rating Scale scores. *British Journal of Psychiatry* **187**, 366–371.
- Lezak MD (1995). *Neuropsychological Assessment. Neuropsychological Assessment*, 3rd ed. Oxford University Press: New York.
- Loong JWK (1990). *The Wisconsin Card Sorting Test (IBM version)*. Wang Neuropsychological Laboratory: San Luis Obispo, CA.
- López-Muñoz F, Álamo C (2011). Neurobiological background for the development of new drugs in schizophrenia. *Clinical Neuropharmacology* **34**, 111–126.
- Mathews SB, Arnold SE, Epperson CN (2014). Hospitalization and cognitive decline: can the nature of the relationship be deciphered? *American Journal of Geriatric Psychiatry* **22**, 465–480.
- Mesholam-Gately RI, Giuliano AJ, Goff KP, Faraone SV, Seidman LJ (2009). Neurocognition in first-episode schizophrenia: a meta-analytic review. *Neuropsychology* **23**, 315–336.
- Moritz S, Birkner C, Kloss M, Jahn H, Hand I, Haasen C, Krausz M (2002). Executive functioning in obsessive-compulsive disorder, unipolar depression, and schizophrenia. *Archives of Clinical Neuropsychology* **17**, 477–483.
- Moritz S, Heeren D, Andresen B, Krausz M (2001). An analysis of the specificity and the syndromal correlates of verbal memory impairments in schizophrenia. *Psychiatry Research* **101**, 23–31.
- Moritz S, Hottenrott B, Jelinek L, Brooks AM, Scheurich A (2012). Effects of obsessive-compulsive symptoms on neuropsychological test performance: complicating an already complicated story. *The Clinical Neuropsychologist* **26**, 31–44.
- Overall JE, Gorham DR (1962). The brief psychiatric rating scale. *Psychological Reports* **10**, 799–812.
- Palmer BW, Dawes SE, Heaton RK (2009). What do we know about neuropsychological aspects of schizophrenia? *Neuropsychology Review* **19**, 365–384.
- Pennington CR, Heim D, Levy AR, Larkin DT (2016). Twenty years of stereotype threat research: a review of psychological mediators. *PLoS ONE* **11**, e0146487.
- Reichenberg A (2010). The assessment of neuropsychological functioning in schizophrenia. *Dialogues in Clinical Neuroscience* **12**, 383–392.
- Reitan RM (1992). *Trail Making Test. Manual of Administration and Scoring*. Reitan Neuropsychology Laboratory: Tucson, AZ, USA.
- Schaefer J, Giangrande E, Weinberger DR, Dickinson D (2013). The global cognitive impairment in schizophrenia: consistent over decades and around the world. *Schizophrenia Research* **150**, 42–50.
- Scheurich A, Fellgiebel A, Schermuly I, Bauer S, Wölfiges R, Müller MJ (2008). Experimental evidence for a motivational origin of cognitive impairment in major depression. *Psychological Medicine* **38**, 237–246.
- Schmidt-Atzert L (2004). *Test d2: Aufmerksamkeits-Belastungs-Test [Test d2: attention-strain-test]*. *Diagnostik von Konzentration und Aufmerksamkeit 6th edn*. Hogrefe: Göttingen.
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* **59**, 22–33.
- Shin NY, Lee TY, Kim E, Kwon JS (2014). Cognitive functioning in obsessive-compulsive disorder: a meta-analysis. *Psychological Medicine* **44**, 1121–1130.
- Spencer SJ, Steele CM, Quinn DM (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* **35**, 4–28.
- Strauss GP, Morra LF, Sullivan SK, Gold JM (2015). The role of low cognitive effort and negative symptoms in neuropsychological impairment in schizophrenia. *Neuropsychology* **29**, 282–291.
- Tombaugh TN (2004). Trail making Test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology* **19**, 203–214.
- Vinogradov S, Fisher M, Warm H, Holland C, Kirshner MA, Pollock BG (2009). The cognitive cost of anticholinergic burden: decreased response to cognitive training in schizophrenia. *American Journal of Psychiatry* **166**, 1055–1062.

- Wechsler D** (2008). *Wechsler Adult Intelligence Scale - Fourth Edition (WAIS-IV)*. Pearson: San Antonio, TX.
- Wilson B, Alderman N, Burgess P, Emslie H, Evans J** (2003). Behavioural assessment of the dysexecutive syndrome (BADS). *Journal of Occupational Psychology, Employment and Disability* **5**, 33–37.
- Wilson RS, Hebert LE, Scherr PA, Dong X, Leurgens SE, Evans DA** (2012). Cognitive decline after hospitalization in a community population of older persons. *Neurology* **78**, 950–956.
- Woodard JL, Axelrod BN** (1987). Wechsler memory scale - revised. *Psychological Assessment* **7**, 445–449. ST-Parsimonious prediction of Wechsler.
- Wykes T, Huddy V, Cellard C, McGurk SR, Czobor P** (2011). A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. *American Journal of Psychiatry* **168**, 472–485.
- Zimmermann P, Fimm B** (1995). *Test for Attentional Performance (TAP)*. PsyTest: Herzogenrath.