LEIBNIZ CENTER FOR RESEARCH IN COMPUTER SCIENCE
TECHNICAL REPORT 2007-03

The distortion of reality perception in schizophrenia patients, as measured in Virtual Reality

Anna Sorkin[1] B.Sc., Daphna Weinshall[1,2] Ph.D., Avi Peled[3,4] M.D.

[1]Interdisciplinary Center for Neural Computation, Hebrew University of Jerusalem, POB 1255, 91904 Israel
[2]School of Computer Science and Engineering, Hebrew University of Jerusalem, 91904, Israel
[3]Institute for Psychiatric Studies, Sha'ar Menashe Mental Health Center, Mobile Post Hefer 38814, Hadera, Israel
[4]Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, 31096 Haifa, Israel

Corresponding author: Daphna Weinshall, daphna@cs.huji.ac.il

Abstract

Background: As a group, schizophrenia patients are impaired on many cognitive tests. Individual patients, however, usually fall within the normal range on many tests, with less than 70% of the patients exhibiting deficiency on each standard test.

Aims: To design an objective test for measuring the distortion in reality perception in schizophrenia patients, and to compare its discriminative power with standard tests.

Methods: 43 schizophrenia patients and 29 healthy controls navigated in a Virtual Reality world and detected incoherencies, like a barking cat or red tree leaves.

Results: Whereas the healthy participants reliably detected incoherencies in the virtual experience, 88% of the patients failed this task. The patient group had specific difficulty in the detection of audio-visual incoherencies; this was significantly correlated with the hallucinations score of the PANSS.

Conclusions: Poor incoherencies detection is a powerful indicator of schizophrenia, more discriminative than most standard cognitive test.

Declaration of interest: None.

1. Introduction

Schizophrenia is a severe mental disorder afflicting 1% of the population world-wide. It is a major economic liability in the western world: in 2002 in the US alone, overall costs linked to schizophrenia were estimated as $62.7 billion (Wu et al., 2005). Even though therapy has achieved considerable progress, schizophrenia still has no cure. To date the pathological mechanisms of this debilitating disorder remain unknown, which reinforces the need in further investigations into the cognitive deficits associated with this disorder.

It is difficult to find any cognitive task that schizophrenia patients perform adequately. The key cognitive dimensions compromised in schizophrenia were recently summarized by NIMH in the MATRICS consensus cognitive battery, including: speed of processing, attention, working memory, verbal learning, visual learning, reasoning and problem solving, and social cognition (MATRICS at http://www.matrics.ucla.edu/provisional-MATRICS-battery.shtml). However, any individual may perform within the normal range on many tasks, and only 9% -67% of schizophrenia patients exhibit impairment in any particular cognitive dimension (Palmer et al., 1997).

Currently, the diagnosis of schizophrenia is routinely established according to the DSM-IV-TR criteria, following the guidelines of the Structured Clinical Interview for DSM-IV Axis I Disorders (First et al., 1995). The severity of schizophrenia is then assessed by the Positive and Negative Syndromes Scale (PANSS) (Kay et al., 1987). Many studies investigated the relationship between cognitive impairment and specific symptomatic sub-groups of the population of schizophrenia patients, such as patients exhibiting either positive or negative symptoms. Though numerous significant correlations were found, they are not always reliably replicated in all studies. Negative symptoms show robust correlations with most cognitive deficit, including: executive function, Wisconsin card sorting test (WCST), trail making test, verbal fluency, working memory, attention, and motor speed (Vasilis et al., 2004). Patients manifesting mainly positive symptoms are considered less impaired. While some studies report the correlation of positive symptoms with working memory (Keefe, 2000), attention (Green and Walker, 1986; Walker and Harvey, 1986; Berman et al., 1997) and verbal memory (Holthausen et al., 1999; Norman et al., 1997), other researches did not find correlation of positive symptoms with working memory or attention (Vasilis et al., 2004; Cameron et al., 2002). Impairment in verbal declarative memory showed correlation with positive symptoms in 8 out of 29 studies (Cirillo and Seidman, 2003).

There is still a need for new cognitive tests that will robustly correlate with positive symptoms, and will discriminate a greater part of the schizophrenia patients. In particular, it seems desirable to develop tests that measure cognitive impairment in complex tasks which involve many different cognitive functions, since the complex nature of the syndrome may manifest itself differently in complex multi-modal tasks. The distortion in reality perception is commonly accepted as a serious manifestation of schizophrenia. The goal of this study was to develop an objective test, that will measure the distortion in reality perception in a complex realistic environment.

Our test design was built upon current leading theoretical perspectives, which portray schizophrenia as a disturbance in integration (Tononi and Edelman, 2000; Friston and Frith, 1995; Peled, 1999). Thus abnormal reality perception may be conceptualized as disruption in integration. For example, auditory hallucinations can occur when speech perception is not constrained by primary visual and auditory inputs, allowing the individual to experience voices of imaginary speakers (David, 2004). To disclose and measure disrupted integration, a powerful measurement tool must be used that challenges the brain in an integrative manner. Virtual Reality (VR) technology appears especially suitable for this purpose: it generates experiences which are complex and multi-modal on the one hand, and fully controllable on the other.

We used a detection paradigm within real-world experiences to measure abnormal reality perception. A subject is required to detect various incoherent events inserted into a normal virtual environment. Everything is possible: a guitar can sound like a trumpet, causing audio-visual incoherency; a passing lane can be pink, and a house can stand on its roof, resulting in visual-visual incoherencies of color and location respectively (see Figure 1). We expect that a well-integrated brain will easily detect these incoherencies, whereas a disturbed, incoherently acting brain will demonstrate poor detection ability.

2. Cognitive Impairment in Schizophrenia

Over a hundred years of research characterized many cognitive deficiencies of schizophrenia patients. As a group, schizophrenia patients are impaired on almost every cognitive task possible. In 2004 the NIMH established the key cognitive dimensions compromised in schizophrenia (MATRICS at http://www.matrics.ucla.edu/provisional-MATRICS-battery.shtml), where speed of processing, memory and attention are considered the most compromised dimensions (Green 2006).

Neurocognitive correlates of schizophrenia symptoms are extensively studied. It is generally agreed that the severity of negative (PANSS) symptoms correlates with most cognitive deficits[6]. The results are less clear cut regarding positive (PANSS) symptoms. For example, in a work (Vasilis et al., 2004) aimed to study the relationship between psychopathology and cognitive functioning, 58 schizophrenia patients were assessed for: executive function, verbal and visual working memory, verbal and visual memory, attention, visuo-spatial ability and speed of processing. Only two measures were found to be correlated with the severity of positive symptoms (mean of a group), including poor performance on semantic verbal fluency (r=0.35, P=0.005) and Trail Making Part A (r=0.43, P=0 .001). No correlation was found between positive symptoms and working memory or attention as reviewed in the literature (Keefe, 2000; Green and Walker, 1986; Walker and Harvey, 1986; Berman et al., 1997).

Other studies give a mixed picture. In one study, positive symptoms were correlated with Digit Span (r=- 0.42, p = 0.02) – a working memory measure, but not correlated with WCST, Trail making A and B, Verbal Fluency and WAIS-R (Berman et al., 1997). In a study dedicated to the relationship between symptoms and working memory, the severity of positive symptoms was found to be uncorrelated with performance on any of the measures (Cameron et al., 2002). In another study, no clear association was found between positive symptom scores and neurocognitive deficits (Voruganti, 1998).

Overall, the extensive review of verbal declarative memory by Cirillo and Seidman (2003) reveals that positive symptoms showed correlation with memory measures in 8 out of 29 studies. However, two main issues complicate the comparison between different studies. First, the positive symptoms group may contain different symptoms in different studies, with some disagreement regarding such measures as depression, disorganization and excitement. Second, many studies test correlation with a group of symptoms, usually summing over all symptoms in a group, and only some look into the correlation with specific symptoms.

Auditory hallucinations are of particular interest. Brebion et al. (2002, 2005, 2006) found a number of measures correlated with auditory hallucinations, including: poor temporal context discrimination (remembering to which of two lists a word belonged), and increased tendency to make false recognition of words not present in the lists or misattributing the items to another source[1]. An association between hallucinations and response bias (reflecting the tendency to make false detections) was also reported in a signal detection paradigms. Bentall and Slade (1985) used a task in which participants were required to detect an acoustic signal randomly presented against a noise background. The authors then compared two groups of schizophrenia patients, who differed in the presence or absence of auditory hallucinations, on the same task. The two

[1] For example, they may confuse the speaker - experimenter or subject, or they may confuse the modality - was an item presented as a picture or a word.

groups were similar in their perceptual sensitivity, but differed in their response bias. Not surprisingly, patients with hallucinations were more willing to believe that the signal was present.

Very few studies examined the diagnostic value of the cognitive tests battery. One possible reason is that any given patient may fall within the normal range in many tasks. The common way to report a cognitive deficiency compares the means of the patient and control populations, measuring the statistical significance of the difference. This procedure blurs out individual differences, i.e. how many patients performed in the normal range, and how many control subjects fell out of the normal range. Some reviews report that less than 40% of schizophrenia patients are impaired (Goldberg and Gold, 1995; Braff et al., 1991), while others state that a fraction of 11% up to 55% of schizophrenia patients perform in the normal range on different tasks (Torrey et al.,1994; Strauss and Silverstein, 1986; Bryson et al., 1993) . It is therefore not clear whether each patient manifests some subset of cognitive impairments, or whether some patients may preserve a completely normal cognitive function.

In an extensive study Palmer et al. (1997) aimed to explore the prevalence of neuropsychological (NP) normal subjects among the schizophrenia population. The authors examined 171 schizophrenia patients and 63 healthy controls using an extensive neuropsychological battery, measuring performance on eight cognitive dimensions: verbal ability, psychomotor skill, abstraction and cognitive flexibility, attention, learning, retention, motor skills and sensory ability. Each dimension was measured by a number of tests. A neuropsychologist rated functioning in each of the eight NP domains described above, using a 9-point scale ranging from 1 (above average) to 9 (severe impairment). A participant was classified as impaired if s/he had impaired score (≥5) on at least two dimensions. Following this procedure, 27.5% of the schizophrenia patients and 85.7% of the controls were classified as NP-normal. 11.1% of the patients and 71.4% of the controls had unimpaired ratings in all 8 dimensions. The proportion of impaired patients in each dimension varied from 9% to 67%.

In light of these disturbing results, it has been argued by Wilk et al. (2005) that although there exists a sub-group of patients that achieves normal scores relatively to the general population, their score may nevertheless be lower than expected from premorbid functioning. In other words, this sub-group might have had a higher than average premorbid score. To test this assumption the authors tested 64 schizophrenia patients and 64 controls individually matched by their Full-Scale IQ score. Now the patient group showed markedly different neuropsychological profile. Specifically, these patients performed worse on memory and speeded visual processing, but showed superior performance on verbal comprehension and perceptual organization. These finding support the hypothesis that cognitive functioning was impaired in these patients relatively to their premorbid level. It’s worth emphasizing that the control group showed a consistent level of performance on all measures, while the patients exhibited a non-uniform pattern, with some measures matching or superior to the controls group, and some inferior.

In summary, although many cognitive deficits were established among schizophrenia patients, the majority of them are correlated with negative symptoms, and each one is only exhibited by a fraction of the patients. Without individual adjustments taking account of one's IQ and possibly other factors, cognitive tests are unable to reliably discriminate schizophrenia patients from the remaining population. Thus there is still a need for cognitive tests that will correlate with positive symptoms, especially with hallucinations, and for tests which will show impairment in a greater part of the patient group.

3. Methods

3.1 Subjects

43 schizophrenia patients were recruited for the study - 23 in-patients from the inpatient population of the Shaa'r Menashe Mental Health Center, and 20 out-patients from the "Hesed veEmuna" hostel in Jerusalem. 29 healthy controls were matched by age, education level and gender to the patient group. Mean age was 32.6 (SD=8.5), with an average of 11.1 (SD=1.8) years of schooling; 19% were females.

All patients had a psychiatric interview with a senior psychiatrist (A.P.). The diagnosis of schizophrenia was established according to the DSM-IV-TR criteria, and symptoms severity was assessed using the Positive and Negative Syndromes Scale (PANSS) (Kay et al., 1987). Exclusion criteria included history of neurological disorders or substance abuse in the previous 3 months.

The study was approved by the Sha'ar Menashe Mental Health Center Review Board, and informed consent was obtained from all participants after the nature of the study was fully explained to them. All subjects volunteered and received payment. They were tested for color blindness by a color naming procedure and anamnesis.

3.2 Experimental Design and Procedure

Subjects sat comfortably in a reclining chair, wearing a Head Mounted Display (HMD) containing the audio and visual devices and a position tracker (Figure 1D). The HMD delivered the virtual reality and created a vivid sense of orientation and presence. Subjects navigated along a predetermined path through a residential neighborhood, shopping centers and a street market (Figure 1). Apart from the incoherencies which were deliberately planted, the virtual environment was designed to resemble the real world as closely as possible. Whenever the path traversed an incoherent event, progress was halted and a one minute timer appeared, during which the subject had to detect the incoherency. Response included marking the whereabouts of the incoherent event by a mouse click, and an accompanying verbal explanation to be recorded. A response was counted as correct only when the subject provided a proper explanation. We gave no examples before the test as guidelines, and no feedback indicating correct or incorrect detection. (A demonstration movie of the virtual world can be found at http://www.cs.huji.ac.il/~daphna/demos.html#incoherencies).

We created three categories of incoherent events: sound (Figure 1C), color (Figure 1A) and location (Figure 1B). The virtual world contained 50 incoherencies: 16 involving color, 18 concerning location and 16 related to sound.

3.3 Data Analysis

Three incoherencies were excluded from the final analysis: two due to the high miss rate (≥25%) among the control subjects, and one due to repeated reports of its being confusing. This resulted in 14 incoherencies of color, 17 - location, 16 – sound, total of 47.

We measured detection rates separately for the sound, color and location categories, as well as the total detection rate and reaction time. We had initially planned to compare the detection rates between the patient and control groups, and investigate the difference between the detection of sound and visual incoherencies, monitoring in particular possible correlations in patients manifesting positive PANSS symptoms. While analyzing the data, we defined and quantified the *gap* parameter, which indicates whether some specific categorical deficiency exists. A *gap* is measured relative to individual performance levels, indicating whether the subject's detection rate in one category differed significantly from the remaining detection rates. Thus a subject could have uniform performance, a *gap* in one category, or a *gap* in 2 categories. For example, if a subject detected color and location incoherencies at a rate of 93% and 88% respectively, and sound at a rate of 25%, he was said to have a gap in the sound category.

For each important parameter, we define its *normal range* as the mean of the control group ±2.5 SD (including roughly 99% of the normal population). We then check for each measurement whether it falls within or outside this range.

4. Results

We analyzed the results in a number of ways. First (Section 4.1), we analyzed the detection rates, which showed a very clear and significant difference between the control group (with close to perfect performance) and the patient group (with typically poor performance). Second (Section 4.2), we analyzed the verbal response of the participants, showing significant difference in the relevance, coherency and length of the answers between the patient and control groups. Third (Section 4.3), we defined and analyzed the *gap* phenomenon, which showed that patients had much larger variability in their responses as compared to the control group. Fourth (Section 4.4), we measured the correlation between the patients' PANSS scores and the measurements obtained in our experiments. Notably, we found a strong correlation between increased hallucinations and poor detection rate in our experiments. Finally (Section 4.5), we analyzed the various types of incoherent events, categorizing them and ranking them according to their discriminability.

4.1 Detection Rates

The histogram of detection rates is shown in Figure 2. The control subjects detected incoherencies very well, with an accuracy level of 96% on average (SD=4) (left panel). In general, the patient group (right panel) differed significantly from the controls. Normal detection rates are shown in red for each category, whereas blue bars indicate the number of subjects that performed below normal. For example, the normal range for total detection rates is 87-100%. The upper plot shows that all but one of the control subjects performed in this range. Among the patients only 6 subjects (red bars) performed in the normal range, whereas 37 subjects (blue bars) had lower detection rates. The patients group exhibited the most difficulty in the sound category: 30 patients performed below the normal range, and 19 had detection rates below 50%, compared to the location category, where only 10 patients detected less than 50% of the incoherencies.

4.2 Analysis of Verbal Response

Detection was only scored as correct when the subject provided a plausible explanation. To determine correctness, a number of external observers, blind to the purpose of the experiment and the assignment to patient vs. control group, analyzed the (recorded) verbal response associated with each incoherency detection. They ranked the answer as correct or incorrect, and provided some additional ranking as explained below

The analysis revealed that about two thirds of the patients experienced some difficulty in explaining the incoherencies, even when they correctly identified the incoherent events. Specifically, the control subjects had on average 1 partial detection, defined as a correct mouse click associated with failure to provide a plausible explanation, with a maximum of 4 partial detections. In contrast, 32 (74%) patients failed to explain 5 or more detected incoherencies, with some patients having more than 20 partial detections.

The biggest difficulty was seen in the sound category, but this may be the result of an apparent attentional bias to sound, which lead subjects to prefer sound emitting objects regardless of the presence (or absence) of incoherency. This is supported by the fact that both the control and patient groups showed highly significant decrease in detection rate of color and location incoherencies when a normal sound event was present in the scene. The control group exhibited 6% decrease (T-test t=3.0430 , df=28, p=0.005), and the patient group – 18% decrease (T-test t =5.5425, df=42, p= 0.000002). We further investigated this assumption by analyzing the data of 23 patients for misses in scenes containing normal sound events, scrutinizing the objects (wrongly) reported as incoherent. We found that a normal sound object is chosen as incoherent on average 3.9 times (SD=2.7), while other objects are chosen with average frequency of only 1.5 times (SD=1); this bias favoring the erroneous selection sound objects is significant (F=21.14, df=51, p=2.93e-05).

We performed a detailed analysis of verbal responses on 15 incoherencies in 10 control subjects and 19 patients. We rated their verbal responses for: (i) distance from target (DT) – measuring the relation between response and target, from 0 – full and correct explanation to 3 – completely unrelated; (ii) length – the number of words in a response, and (iii) the number of unrelated topics in the response. The patient group deviated more often from the target stimulus: average DT = 1, as compared to the control group with average DT = 0.17 (ANOVA p=3.3207 e-004, df=27, F=16 .88). The patients also gave longer answers: average length of 15 words vs. 9 in the control group.

4.3 *Gap* Phenomenon and Various Divisions of the Patient Group

The control group showed similar detection rates in all three categories (Figure 3A). The patient group, on the other hand, could be divided into two major sub-groups based on the similarity in detection rates. (1) The *uniform* group – patients whose detection rates in all three categories were similar. (2) *Gap* – the group of patients having specific difficulty in one or two categories. A patient was defined as having a specific impairment in one category – or *gap* – if this category score was significantly below his/her best category (a significant difference is a difference exceeding the mean±2.5SD of the control group). The *uniform* group could be further divided into: i) *uniform normal*: patients performing at normal levels (N=5 subjects, Figure 3B); ii) *uniform fair*: patients with good detection rates (50-87%) but below the normal range (N=10 subjects, Figure 3C); and finally iii) *uniform poor*: patients with poor uniform performance below 50% (N=8 subjects, Figure 3D). Almost half of the patients (the *gap* group) had specific difficultly in one or two categories. 16 patients (37%) had a specific difficulty in detecting audio-visual incoherencies: 7 patients had difficulty in the sound category only (Figure 3E), 7 patients had difficulty in the sound and color categories as compared to the location category (Figure 3F), and 2 patients had difficulty in the sound and location categories. Only 4 patients had other specific difficulties.

4.4 Symptom Analysis

4.4.1 Symptoms across different patient subgroups

Positive symptom scores as measured by PANSS increased across the four patient subgroups: *uniform normal, uniform fair, uniform poor,* and *gap* (Figure 4A). The *uniform normal* group differed significantly from the other three on the ‘hallucinations’ score, as well as the ‘delusions’ score (with a significant difference with the *gap* group). Negative scores showed greater similarity among the four groups, except ‘difficulty in abstract thinking’ where a significant difference was found between the *uniform normal* and *uniform fair* groups and the *uniform poor* and *gap* groups (Figure 4B).

4.4.2 Correlations with symptoms

We found a number of significant correlations (Spearman’s $r \geq 0.3$, $t \geq 2.02$, $df=41$, $p<0.05$) between detection rates and the PANSS scores in the patient group: i) The ‘hallucinations’ score was correlated with low total and sound detection rates. ii) ‘Difficulty in abstract thinking’ showed a correlation with low total, sound and color detection rates (two last correlations: Spearman’s $r \geq 0.3885$, $t \geq 2.7$, $df=41$, $p<0.01$). In addition, reaction time showed a negative correlation with age.

4.4.3 Comparative performance among patient subgroups defined by symptoms

We divided the patients into three groups based on their PANSS scores: i) dominant positive symptoms (N=9); ii) dominant negative symptoms (N=21); and iii) combined group (N=10); 2 patients had no symptoms. The Positive group showed significantly lower detection rates in all categories as compared to the two other groups (Figure 4C).

Surprisingly, the combined group performed similarly to the negative group; i.e., had significantly better detection rates than the positive group in all categories, while maintaining a similar average positive score to the positive group.

In addition, the out-patients performed better than in-patients: i) Total detection rates were on average 10% better; ii) only 2 out-patients had a total detection rate below 50% as compared to 9 in-patients; iii) 4 out of the 5 patients who performed in the normal range were out-patients.

4.5 Analysis of Incoherencies

To evaluate which incoherencies were most successful in discriminating between the control and the patient groups, we used a measure of Mutual Information (MI). Each incoherency is given a high MI score if success or failure to detect it correlates highly with one group alone (control or patients). For example, an incoherency that is only missed by patients is a good discriminator between the groups. An incoherency that is equally detected or missed by the control and patient groups is a poor discriminator.

The 10 most discriminating incoherencies included 6 from the sound category, and 2 from each of the color and location categories. For the patient group these incoherencies were more difficult to detect than the remaining 40, while for the controls they did not present any special difficulty. Examples include: adults laughing like babies, reversed traffic-light colors, floor washing accompanied by the sound of toilet flushing, airplane accompanied by bombing sounds, a bouncing ball sounding like a bell, a blue cola machine, reverse writing on a street sign, and bus making an elephant sound.

The 10 least discriminating incoherencies contained 6 from the location category, and 2 from each of the sound and color categories. These incoherencies were equally easy (or hard) to detect for the patient and control groups. This set of incoherencies included: a dog serving customers, a giraffe shopping, a hydrant in the middle of the road, purple bananas, a chair on the roof, ambulance making an ice-cream-truck melody, a red cloud, a barking cat, a mannequin with a lion-head, and two cows in a bus station.

A closer look at the sound incoherencies revealed that incoherent sounds could be further classified in terms of their relationship to objects: i) same category incoherency, such as a barking cat where one animal's voice is replaced by another animal's voice (animal-animal), or a car making train sounds (vehicle-vehicle replacement); ii) different category, such as a construction truck making gun fire sounds ; and finally iii) same object, when the sound is correct but the circumstances are wrong, like adults laughing as babies, floor washing accompanied by toilet flushing sounds, and a civilian plane making bombing sounds. The last group was the most difficult for the patient group to detect - less than 50% of the patients detected these events, as compared to 92% of the controls.

5. Discrimination Procedure

How well can performance on an incoherencies detection task discriminate between the control and schizophrenia populations? Can we do better than the battery of cognitive tests examined by Palmer et al. (1997), which showed only partial discrimination ability?

We designed a discrimination procedure based on 5 parameters: the four detection scores (total, color, location and sound) and the presence of a *gap*. Thus each subject having 2 or more scores (out of 5) below the normal range was classified as a 'patient', otherwise s/he was defined as 'normal'. This procedure yielded 89% correct classification, with 3.4% false alarms (one healthy subject classified as a patient), and 16.3% misses (7 patients classified as normal), see Table 1A. Next, we removed the 10 least discriminating incoherencies as defined by the MI analysis, in order to improve prediction accuracy to 91.6% (1 control and 5 patients misclassified).

We used a cross-validation paradigm to check the generality of our results and to avoid the danger of over-fitting. Specifically, we divided the subject population into two balanced groups: one with 35 subjects (14 controls and 21 patients), and one with 37 subjects (15 controls and 22 patients). We then calculated the MI measures and the normal ranges using the first group only, and evaluated the discrimination procedure on both groups separately (see Table 1B).

Clearly prediction accuracy is similar in both groups. In addition, when removing the 10 least discriminating incoherencies as calculated based on the first group, we obtained a similar improvement in classification in both groups. This confirms the generality of our results as regards discrimination between the schizophrenia patients and normal populations.

As already mentioned, incoherency detection was counted as correct only when accompanied by an appropriate verbal explanation, leading to observer-dependent variability. We therefore repeated the entire analysis above based on partial detections alone; namely, detection was scored as correct whenever the incoherent object was selected. Despite major improvement in detection rates among the patients, the accuracy of the classification procedure decreased only moderately, correctly classifying 77% as compared to 88% of the patients, and 84% as compared to 92% of the control subjects.

The biggest difference was found in the sound category, where the number of patients failing to detect 50% or more of the incoherencies decreased from 44% to 27%, and the gap group now containing subjects with specific difficulty in color rather than sound. Probably because sound events attract immediate attention regardless of any incoherency (as discussed above in Section 4.2). The analysis of partial detections and the attention bias to sound objects led us to conclude that correct incoherencies detections cannot be used in isolation, and should be accompanied by proper verbal explanation.

6. Comparison with Standard Cognitive Tests

Our assessment design is highly discriminative as compared to most cognitive assessment tests, with 88% of the patients exhibiting impairment in the task; other cognitive tests discriminate correctly only 9-67% of the patients (who perform below the normal range) (Palmer et al., 1997).

To evaluate our test's strength we use a standard measure of effect size - Cohen's d (1988), which estimates the degree to which the phenomenon is present in the population. Specifically, size effect measures the difference between the patient and control means on a variable of interest, calibrated by pooled standard deviation units. In our experiment we obtain an effect size for total detection rate of 1.86, which is a very large effect. For comparison, in a meta-analysis of 204 cognitive studies, Heinrichs and Zakzanis (1997) summarized the mean effect size for different cognitive tests. The biggest effect size was found for global verbal memory and equaled 1.41 (SD=0.59). Other standard tests show smaller effect size. For example, Continuous performance test - 1.16 (SD=0.49), Wisconsin card sorting test - 0.88 (SD=0.41), and Stroop - 1.11 (SD=0.49).

In addition, as the patient's hallucinations become more severe, the detection of audio-visual incoherencies gets worse. This fact suggests that hallucinating patients may suffer from a specific disturbance in audio-visual integration. This may be particularly useful as only few cognitive tests showed any correlation with the presence of hallucinations (Brebion et al., 2002, 2005, 2006; Bentall and Slade, 1985).

The analysis of individual incoherencies revealed that some incoherencies discriminate between the control and patient populations better than others. Thus auditory events proved to be the most effective. Interestingly, we observed that most effective were events involving auditory stimuli, where the object and sound matched overall but were used under the wrong circumstances, as in adults who appear to be laughing but sound like babies laughing.

7. Summary and Discussion

In this study we showed that schizophrenia patients can be readily differentiated from the normal population based on their performance in the Incoherencies Detection Task. Thus this task is a powerful test of schizophrenia deficits, where poor performance correlates with the presence of hallucinations. The task has additional advantages: it is short - taking only half hour, and it can be self-administrated requiring only minimal non-professional assistance. The incoherencies set may be further improved to shorten the duration of the test, and to increase the discriminability of the patient population. The results should also be confirmed with additional comparison groups, consisting of patients with different mental disorders.

In a previous study Sorkin et al. (2006) showed how a virtual environment can be designed to elucidate disturbances of working memory and learning in schizophrenia patients. The measures collected during the working memory task correctly identified 85% of the patients and all the controls. Thus both tests show high discriminability of the schizophrenia and control populations, better than almost any other standard test. We believe that two factors contributed to the success of these tests: (i) conceptualizing schizophrenia as a disturbance in integration and designing tests that will address possible integration deficits, and (ii) using virtual reality as an experimental tool that challenges the brain in an interactive multi-modal way.

Today, when the diagnostic approach to mental disorders in general, and to schizophrenia in particular, is under major discussion (Kendell and Jablensky, 2003; Frances and Egger 1999) and NIMH calls for the development of new approaches (Kupfer et al., 2005), the neurocognitive testing can provide the desired alternative. Based on the evaluation of eight cognitive dimensions, Palmer et al. (1997) predicted correctly 72.5% of the patients and 85.7% of controls. By developing additional cognitive tests addressed at integration, the diagnostic power of the tests can be increased. Thus describing a patient by a performance profile, containing measurements taken during cognitive tests rather than symptoms, offers benefits to both the patient and the treating psychiatrist: the measures are objective, each patient receives a unique characterization, and cognitive deficiencies are readily related to neuro-scientific knowledge. Given the current state of affairs, it seems that many more experiments are required before a successful diagnostic profile of schizophrenia can be constructed.

Acknowledgments

The authors thank the staff of the Hesed and Emuna hostel in Jerusalem and its director Hannah Rosenthal for their help and encouragement.

References

1. Bentall R.P. and P.D. Slade, 1985. Reality testing and auditory hallucinations: a signal detection analysis. British Journal of Clinical Psychology 24 159–169.
2. Berman I, Viegner B, Merson A, Allan E, Pappas D,Green AI., 1997. Differential relationships between positive and negative symptoms and neuropsychological deficits in schizophrenia. Schizophr Res 25:1-10.
3. Braff DL, Heaton R, Kuck J, Cullum M, Moranville J, Grant I, Zisook S., 1991. The generalized pattern of neuropsychological deficits in outpatients with chronic schizophrenia with heterogeneous Wisconsin Card Sorting Test results. Arch Gen Psychiatry 48:891–898.
4. Brebion, G., Gorman, J., Amador, X., Malaspina, D., & Sharif, Z., 2002. Source monitoring impairments in schizophrenia: Characterisation and associations with positive and negative symptomatology. Psychiatry Research, 112, 27–39.
5. Brebion G, David AS, Jones H, Pilowsky LS., 2005. Hallucinations, negative symptoms, and response bias in a verbal recognition task in schizophrenia. Neuropsychology. Sep;19(5):612-7.
6. Brebion G, David AS, Jones HM, Ohlsen R, Pilowsky LS., 2006. Temporal context discrimination in patients with schizophrenia: Associations with auditory hallucinations and negative symptoms. Neuropsychologia. Sep 20.
7. Bryson, G. J., Silverstein, M. L., Nathan, A., & Stephen, L.,1993. Differential rate of neuropsychological dysfunction in psychiatric disorders: Comparison between alstead-Reitan and Luria-Nebraska batteries. Perceptual and Motor Skills, 76, 305-306.

8. Cameron AM, Oram J, Geffen GM, Kavanagh DJ, McGrath JJ, Geffen LB., 2002. Working memory correlates of three symptom clusters in schizophrenia. Psychiatry Res. 15;110(1):49-61.
9. Cirillo, M. A., & Seidman, L. J., 2003. Verbal declarative memory dysfunction in schizophrenia: From clinical assessment to genetics and brain mechanisms. Neuropsychology Review, 13, 43–77.
10. Cohen, J. 1988. Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press.
11. David AS., 2004. The cognitive neuropsychiatry of auditory verbal hallucinations: an overview. Cognit Neuropsychiatry. Feb-May;9(1-2):107-23.
12. First, M., Spitzer, R.L., Gibbon, M. and Williams, J.B.W., 1995, SCID (DSM-IV) Structured Clinical Interview for Axis I DSM-IV Disorders - Patient Edition (SCID-I/P). Biometrics Research Department, New York State Psychiatric Institute: New York.
13. Frances AJ, Egger HI, 1999. Whither psychiatric diagnosis Aug NZJ Psychiatry. 33:161-165.
14. Friston KJ, Frith CD., 1995. Schizophrenia: a disconnection syndrome? Clin Neurosci 3(2):89-97.
15. Goldberg, T.E. and Gold, J.M., 1995. Neurocognitive functioning in patients with schizophrenia. In: Bloom, F.E. and Kupfer, D.J., Editors, 1995. Psychopharmacology: The Fourth Generation of Progress, Raven Press, New York, pp. 1245–1257.
16. Green M, Walker E., 1986. Attentional performance in positive and negative symptom schizophrenia. J Nerv Ment Dis 174:208-213.

17. Green Michel F., 2006. “Cognitive Impairment and Functional Outcome in Schizophrenia and Bipolar Disorder” . J Clin Psychology 67 (suppl 9):3-8.

18. Heinrichs W. and Zakzanis K.K.,1998. Neurocognitive Deficit in Schizophrenia:A Quantitative Review of the Evidence. Neuropsychology, Vol. 12, No. 3,426-445.

19. Holthausen EAE, Wiersma D, Knegtering RH, Van den Bosch RJ., 1999. Psychopathology and cognition in schizophrenia spectrum disorders: the role of depressive symptoms. Schizophr Res 39:65-71.

20. Kay, S.R., Fiszbein, A. and Opler, L.A., 1987, The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261-276.

21. Keefe RSE. 2000. Working memory dysfunction and its relevance to schizophrenia. In: Sharma T, Harvey P (eds). Cognition in Schizophrenia: Impairments, Importance and Treatment Strategies. New York, NY: Oxford University Press, 16- 50.

22. Kendell R and Jablensky A., 2003. Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry. Jan;160(1):4-12.

23. Kupfer D. J., First B.B., Regier D. A, 2005. A Research Agenda for DSM-V. Published by the American Psychiatric Association.

24. MATRICS: Measurement and Treatment Research to Improve Cognition in Schizophrenia. 2004. MATRICS: Provisional Consensus Cognitive Battery. Available at: http://www.matrics.ucla.edu/provisional-MATRICS-battery.shtml. Accessed Oct 22, 2006.

25. Norman RMG, Malla AK, Morrison-Stewart SL, Helmes E, Willianson PC, Thomas J, Cortese L., 1997. Neuropsychological correlates of syndromes in schizophrenia. Br J Psychiatry 170:134-139.

26. Palmer, B.W., Heaton, R.K., Paulsen, J.S., Kuck, J., Braff, D., Harris, M.J., Zisook, S. and Jeste, D.V., 1997. Is it possible to be schizophrenic yet neuropsychologically normal?. Neuropsychology 11, pp. 437–446.
27. Peled A., 1999. Multiple contraint organization in the brain: a theory for schizophrenia. Brain Res Bull 49(4):245-50.
28. Sorkin A, Weinshall D, Modai I, Peled A., 2006. Improving the accuracy of the diagnosis of schizophrenia by means of virtual reality. Am J Psychiatry. Mar;163(3):512-20.
29. Strauss, B. S., & Silverstein, M. L., 1986. Luria-Nebraska measures in neuropsychologically nonimpaired schizophrenics: A comparison with normal subjects. International Journal of Clinical Neuropsychology, 8, 35-38.
30. Tononi G, Edelman GM., 2000. Schizophrenia and the mechanisms of conscious integration. Brain Res Brain Res Rev. 31(2-3):391-400.
31. Torrey, E. E, Bowler, A. E., Taylor, E. H., & Gottesman, I. I., 1994. Schizophrenia and manic-depressive disorder. New York: Basic Books.
32. Vasilis P. Bozikas, Mary H. Kosmidis, Konstantina Kioperlidou, Athanasios Karavatos, 2004. “Relationship Between Psychopathology and Cognitive Functioning in Schizophrenia” Comprehensive Psychiatry, 45 (5) 392-400.
33. Voruganti LN, Heslegrave RJ, Awad AG., 1998. Neurocognitive correlates of positive and negative syndromes in schizophrenia. Can J Psychiatry. Oct;43(8):854.
34. Walker E, Harvey P., 1986. Positive and negative symptoms in schizophrenia: attentional performance correlates. Psychopathology 19:294-302.

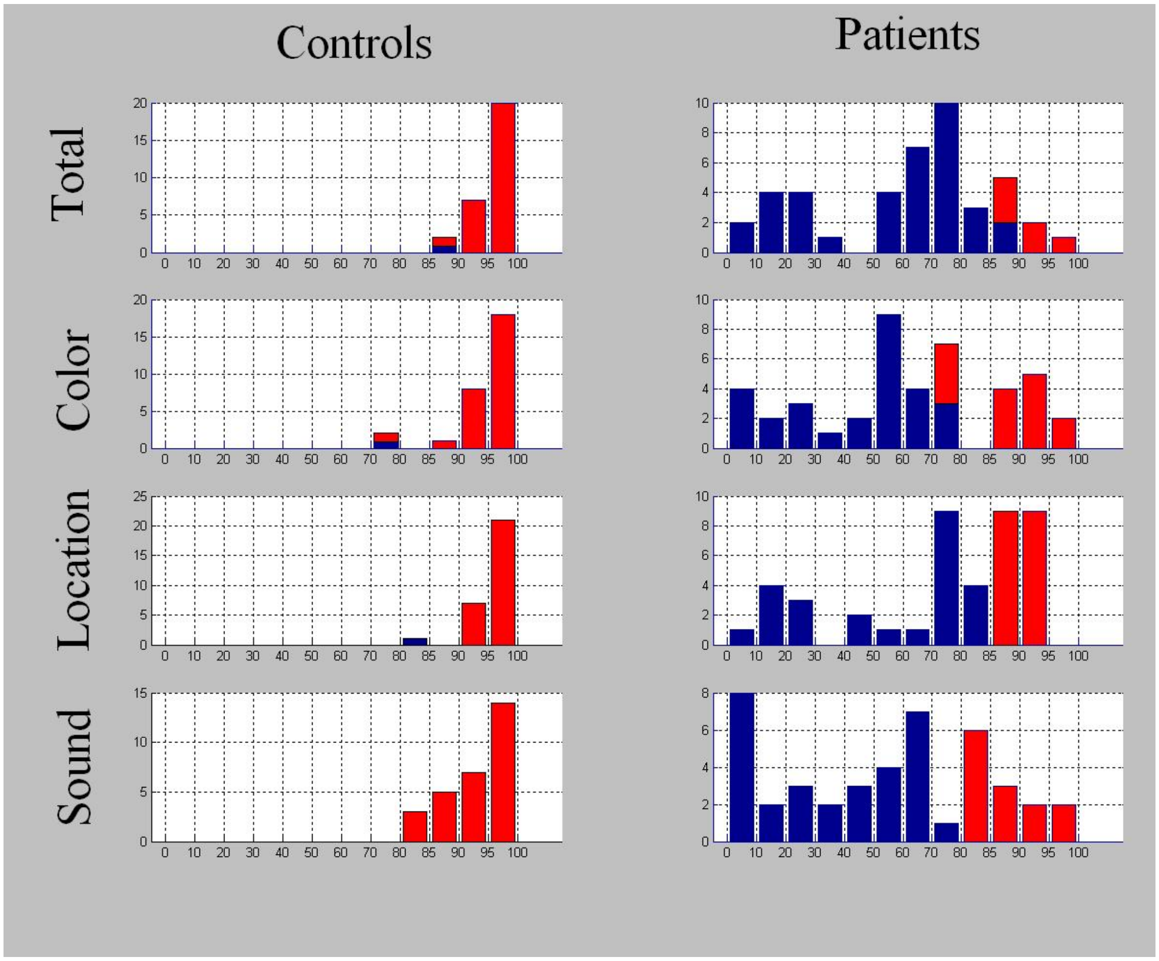
35. Wilk CM, Gold JM, McMahon RP, Humber K, Iannone VN, Buchanan RW., 2005. No, it is not possible to be schizophrenic yet neuropsychologically normal. Neuropsychology Nov;19(6):778-86.

36. Wu EQ, Birnbaum HG, Shi L, Ball DE, Kessler RC, Moulis M, Aggarwal J., 2005. The economic burden of schizophrenia in the United States in 2002. J Clin Psychiatry. Sep;66(9):1122-9.

Figure 1. Examples from the virtual world used in the experiment.

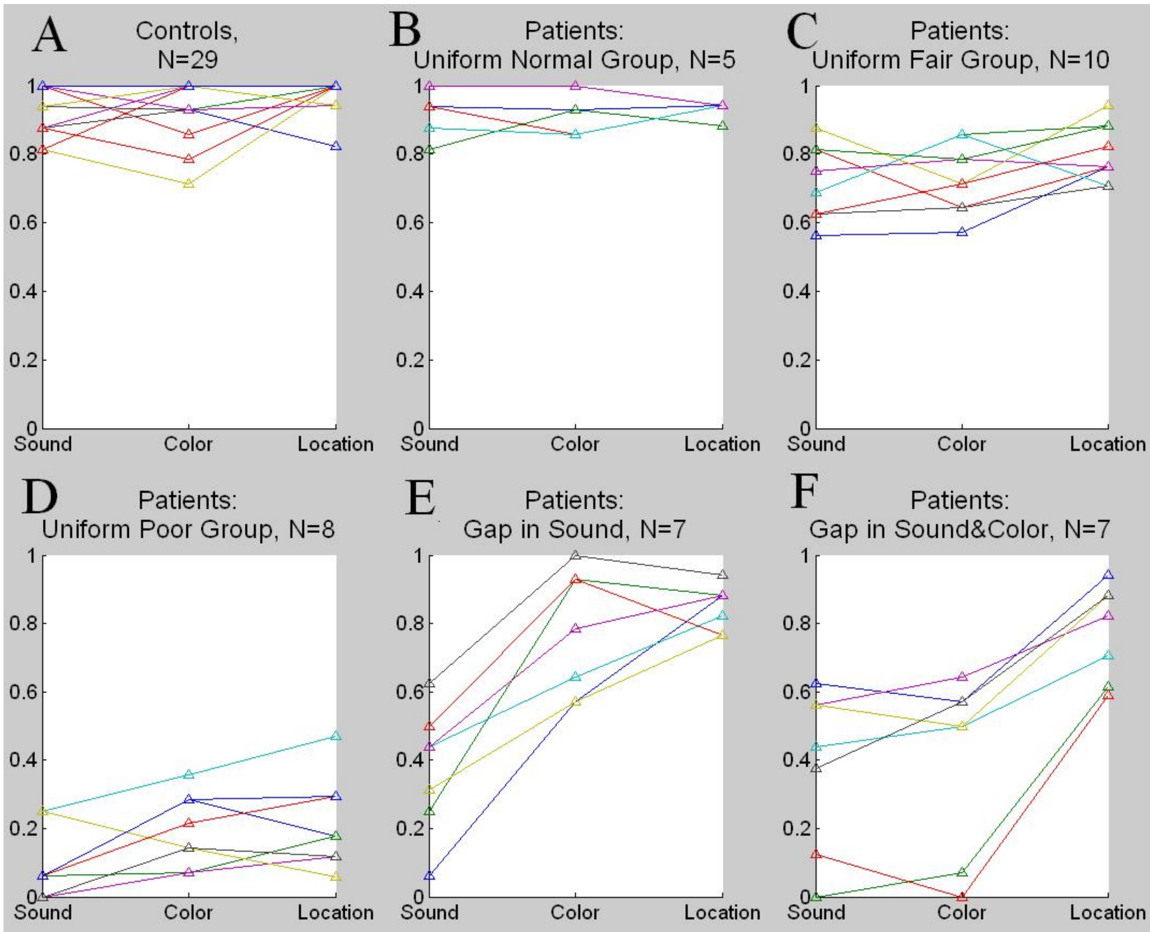
A. incoherent color; **B**. incoherent location; **C**. incoherent sound: a guitar emitting trumpet sounds, and an ambulance sounding like an ice-cream truck.

Figure 2. **Histogram of detection rates among the control and patient groups.**



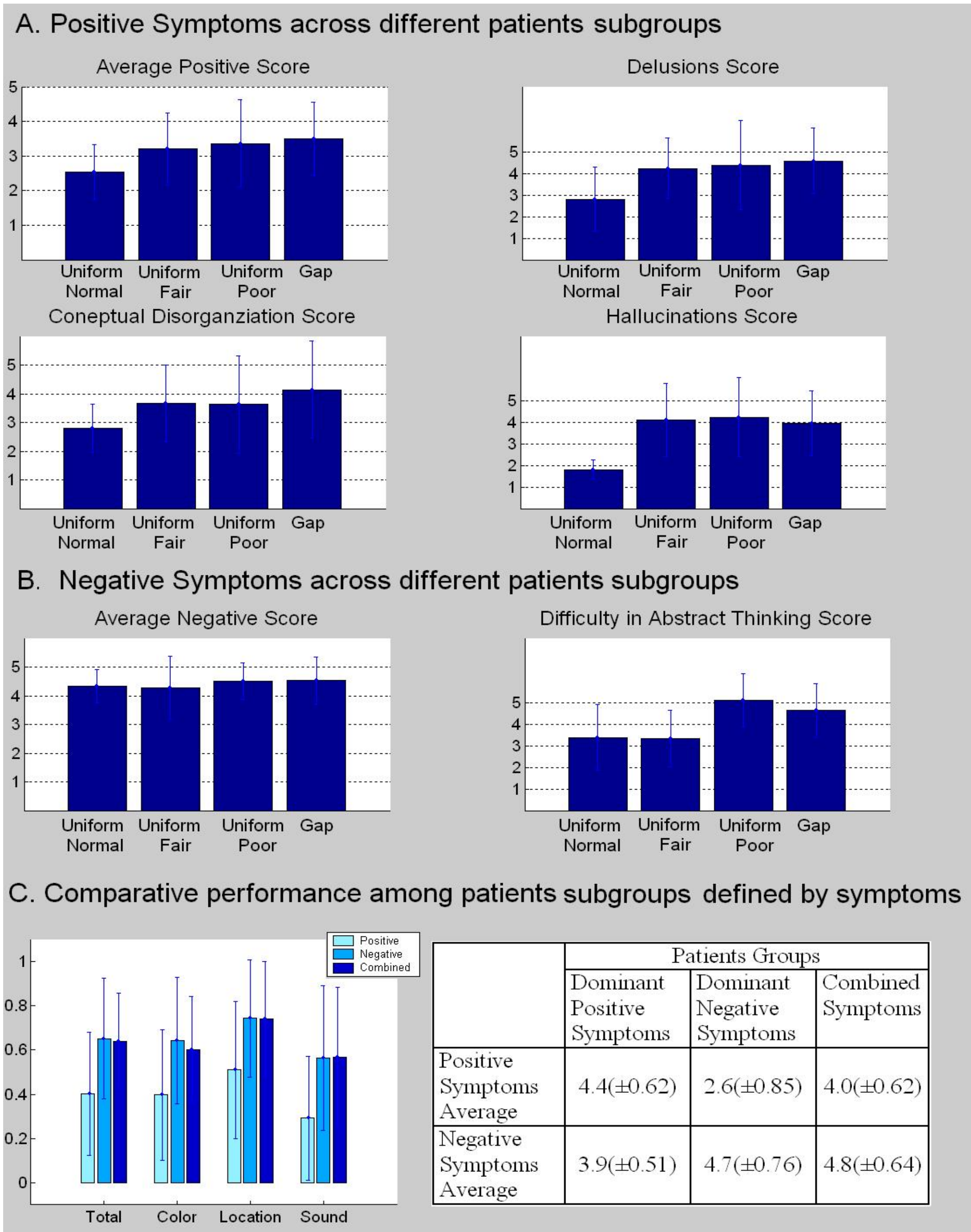
Horizontal axis represents detection rate, vertical axis shows the number of subjects obtaining each score. The red bars indicate performance in the normal range, and the blue bars – performance beyond the normal range.

Figure 3. Individual detection rates of the control and patient groups.



A. Controls. **B-E.** The patients' subgroups. **B**. Uniform normal; **C**. Uniform fair; **D**. Uniform poor; **E.** Gap in the sound category; **F.** Gap in the sound and color categories.

Figure 4. A&B. Selected PANSS scores for four patient subgroups. **C.** Comparative performance among patients subgroups defined by symptoms: dominant positive symptoms; dominant negative symptoms; and combined symptoms. Left panel shows detection rates and right panel shows symptom statistics for each group.



	Patients Groups		
	Dominant Positive Symptoms	Dominant Negative Symptoms	Combined Symptoms
Positive Symptoms Average	4.4(±0.62)	2.6(±0.85)	4.0(±0.62)
Negative Symptoms Average	3.9(±0.51)	4.7(±0.76)	4.8(±0.64)

Table 1. Improvement in correct prediction rates after removing the 10 least discriminating incoherencies.

A. Analysis performed on all subjects. **B**. Cross-validation test: removal of incoherencies was calculated using only half the subjects – the first group.

	A	All Subjects		**B**	First group		Second group	
		All features	Removing 10 easy		All features	Removing 10 easy	All features	Removing 10 easy
Controls		96.5%	96.5%		93%	93%	100%	100%
Patients		84%	88%		81%	90.5%	82%	86.4%
Total		89%	91.6%		86%	91%	89%	92%