
BRIEF REPORT

Reliable Change Indexes for Memory Performance in Schizophrenia as a Means to Determine Drug-Induced Cognitive Decline

Steffen Moritz

University of Hamburg, Hamburg, Germany

Grant L. Iverson and Todd S. Woodward

University of British Columbia and Riverview Hospital, Port Coquitlam, British Columbia, Canada

Some psychotropic medications frequently prescribed in schizophrenia, especially agents that contain anticholinergic properties, have adverse effects on learning and memory. Given the impact of memory on both functional and symptomatic outcome in schizophrenia, it is important to determine if patients suffer from medication-related memory impairment. This study provides a statistical methodology for identifying medication-related changes in memory functioning. Reliable change estimates were determined for Rey Auditory Verbal Learning Test performance in 38 inpatients with schizophrenia. When an 80% confidence criterion was applied, improvement or decline of 12 words in the total score (sum score: Trials 1–5) cannot be attributed to measurement error. These findings should help clinicians better understand cognitive side effects and facilitate decision making regarding changes to individual patient's psychopharmacotherapy.

Key words: schizophrenia, memory, neuroleptics, antipsychotics, reliable change

A large body of evidence suggests that some psychotropic substances commonly prescribed in the treatment of schizophrenia may have significant adverse effects on learning and memory (Buffett-Jarrott & Stewart, 2002; Medalia, Gold, & Merriam, 1988; Moritz, 2002). Cognitive deficits are most common with medications containing anticholinergic properties, such as anti-parkinsonian agents (e.g., bztropine; Saint-Cyr, Taylor, & Lang, 1993; Spohn & Strauss, 1989; Sweeney, Keilp, Haas, & Weiden, 1991). These medications frequently are administered to ameliorate the neurological side effects of conventional neuroleptic

drugs like haloperidol. In addition, memory decline has occasionally been associated with clozapine (Carter, Thrasher, & Thornicroft, 1994). However, most studies have revealed beneficial effects of clozapine on memory performance (Grace et al., 1996; Lee, Jayathilake, & Meltzer, 1999; see also Moritz, 2002, for a review). In addition, other so-called atypical neuroleptics such as risperidone and olanzapine have been found to improve some aspects of memory performance (Meltzer & McGurk, 1999). Because memory deficits in schizophrenia have a strong negative impact on the functional outcome, such as independence and social relations (Green, 1996), and have also been found to negatively affect symptomatic outcome (Moritz et al., 2000; Moritz, Perro, Woodward, & Krausz, 2002), detection and treatment of such dysfunctions is important.

The aim of this study is to help clinicians determine if a patient with schizophrenia has suffered from drug-induced memory decline or has benefited from a

The authors thank Richard Hughes for his assistance with manuscript preparation.

Requests for reprints should be sent to Steffen Moritz, Universitätsklinikum Hamburg–Eppendorf, Klinik für Psychiatrie und Psychotherapie, Martinistraße 52; D-20246 Hamburg, Germany. E-mail: moritz@uke.uni-hamburg.de

neuropsychological or psychopharmacological intervention by providing statistical parameters for determining meaningful change based on objective memory assessment. Such an inference should not be based solely on clinical judgment or patients' self-reports; clinical judgments can differ from objective neurocognitive results. Furthermore, patients' self-reports might be biased by influences other than core neurocognitive complaints.

To meet our objective, we used a modification of the widely used reliable change (RC) methodology initially proposed by Jacobson and Truax (1991). The RC methodology has been featured prominently in the psychotherapy literature (e.g., Hageman & Arrindell, 1993; Hsu, 1989; Jacobson & Revenstorf, 1988; Jacobson, Roberts, Berns, & McGlinchey, 1999; Jacobson & Truax, 1991; Ogles, Lambert, & Masters, 1996; Speer, 1992; Speer & Greenbaum, 1995). More recently, this methodology has been used in clinical neuropsychology (e.g., Chelune, Naugle, Luders, Sedlak, & Awad, 1993; Heaton et al., 2001; G. Iverson, 1999; G. I. Iverson 2001; G. L. Iverson, 1998; Temkin, Heaton, Grant, & Dikmen, 2000) and sports neuropsychology (Barr & McCrea, 2001; Hinton-Bayre, Geffen, Geffen, McFarland, & Friis, 1999). This methodology provides indexes of change that are unlikely to occur as a result of chance. For this study, the Rey Auditory Verbal Learning Test (RAVLT) was used, a widely used memory test to measure learning difficulties (Lezak, 1995).

and Psychotherapy) were initially recruited. Diagnoses were confirmed by at least two clinicians following a semistructured interview. Exclusion criteria were substantial drug or alcohol abuse and substantial neurological impairment (e.g., stroke, concussion, epilepsy), or any medical condition incompatible with a diagnosis of schizophrenia, such as Alzheimer's dementia. Patients with comorbid diagnoses such as depression or anxiety were not excluded, because we wanted a representative sample of patients with schizophrenia. Patients underwent initial neuropsychological assessment 2 weeks after being stabilized on atypical neuroleptics following a short wash-out phase (3–7 days). Patients receiving anti-parkinsonian medication ($n = 1$) or clozapine ($n = 3$) were excluded. Participants were reassessed after 2 weeks. Memory data on both occasions was available for 38 patients. A control group composed of 31 healthy participants was recruited mainly from hospital staff or members of the German armed forces. Healthy participants were free of any psychiatric and neurological disturbances. Demographic characteristics of the samples are provided in Table 1. The control group was used to (a) demonstrate that patients with schizophrenia typically perform more poorly on the test of learning and memory, (b) examine the test-retest reliability of the German alternate forms of the test, and (c) determine whether there are practice effects on the alternate forms of the test.

Methods

Participants

Forty-nine inpatients treated for schizophrenia at the University Hospital of Hamburg (Hospital for Psychiatry

Measure

Version A of the RAVLT was administered at baseline and version B 2 weeks later. Both lists were a direct German translation (Heubrock, 1992) of the English test (Lezak, 1995). Participants were instructed that a list of

Table 1. Demographic Characteristics and Memory Performance of the Samples

Variable	Healthy Participants ^a		Schizophrenia ^b		Statistics
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Sex (male/female)	20/11		26/12		$\chi^2 = .12; p > .7$
Age	33.03	11.26	32.45	9.98	$t(67) = .23; p > .8$
Education	11.77	1.45	11.47	1.48	$t(67) = .84; p > .3$
Premorbid Functioning (MWT-B)	31.06	3.44	28.44	5.77	$t(67) = 2.22; p = .029$
Hospitalizations	—	—	3.14	4.63	—
Length of Illness	—	—	5.75	6.91	—
RAVLT Sum 1–5 (baseline)	59.16	6.09	46.34	12.26	$t(67) = 5.65; p < .001$
RAVLT Sum 1–5 (retest)	57.84	8.51	45.66	12.06	$t(67) = 5.91; p < .001$

Note: MWT-B = The Mehrfachwahl-Wortschatztest (Multiple Choice Vocabulary Test) assesses passive knowledge of German words; RAVLT = Rey Auditory Verbal Learning Test.

^a $n = 31$. ^b $n = 38$.

words would be read to them that they had to recall immediately afterwards. The experimenter then read the words slowly in a monotone voice. The list was presented five times. Subsequently, a new list was presented (Interference List), and participants were asked to recall the words from that list. Participants were then asked to recall the words from the original list (i.e., short-delay free recall) without prior presentation. After a 20-min delay, participants were asked to recall the original list again. The total number of correctly recalled items from the first five trials was the primary dependent variable.

Data Analyses

The RC methodology allows the clinician to minimize the impact of measurement error on test interpretation. To assess clinically significant improvement, it is important that the change score be statistically reliable. However, a statistically reliable change does not necessarily imply a clinically meaningful change. For example, if a patient with schizophrenia demonstrated significant memory problems that appeared to be medication related, and then obtained a memory test score that showed statistically reliable improvement after a change in medications (yet the score was still very low). This might not indicate clinically meaningful improvement if the score remains in the impaired range and real-world memory problems persist.

The standard error of the difference (S_{diff}) can be used to create a confidence interval (CI; a prediction interval in the statistical literature) for test–retest difference score. Essentially, this confidence interval represents the probable range of measurement error for the distribution of difference scores. The formula for calculating the S_{diff} is as follows:

$$SEM_1 = SD\sqrt{1 - r_{12}}$$

Standard deviation from time 1 multiplied by the square root of 1 minus the test–retest coefficient.

$$SEM_2 = SD\sqrt{1 - r_{12}}$$

Standard deviation from time 2 multiplied by the square root of 1 minus the test–retest coefficient.

$$S_{\text{diff}} = \sqrt{SEM_1^2 + SEM_2^2}$$

Square root of the sum of the squared SEMs for each testing occasion.

Note that the formula for the S_{diff} includes the SEM for time 1 and time 2. The formula, as originally presented in Jacobson and Truax (1991), was incorrect. Unfortunately, nearly every study reported in the psychotherapy and neuropsychology literature uses the incorrect formula, including the most recent studies (e.g., Barr & McCrea, 2001; Temkin et al., 2000). This formula should be considered an estimated S_{diff} (estimated $S_{\text{diff}} = \sqrt{SEM_1^2}$; Hageman & Arrindell, 1993; G. I. Iverson, 2001; G. L. Iverson, 1998), not the actual S_{diff} . This formula represents an estimated standard error of difference because the SEM for time 1 is prorated instead of using the SEM for time 2. This formula can be used out of necessity when participants have not undergone follow-up testing (e.g., G. I. Iverson, 2001; G. L. Iverson, 1998). Under these circumstances, the S_{diff} is estimated from a single testing session.

The confidence interval for the test–retest difference score is obtained by multiplying the S_{diff} by a value from the z distribution. Multiplying by a value of 1.64, for example, results in a change in score in either direction that would be unlikely to occur by chance ($p < .05$ in each tail). Multiplying by a value of 1.28 forms a .80 CI ($p < .10$ in each tail).

One of the biggest controversies with the RC methodology has been how to handle regression to the mean. Speer (1992) attempted to improve the reliable change interval reported by Jacobson, Follette, and Revenstorf's (1984) by correcting for the effects of regression to the mean. Hsu (1989, 1999) and Speer (1992) proposed alternate RC formulas to correct for the effects of regression to the mean. Further refinements to the methodology are proposed by Hageman and Arrindell (1993, 1999a, 1999b). The issue is far from resolved. We chose to use the method that does not correct for regression, but does correct for practice (Chelune et al., 1993; G. L. Iverson & Green, 2001) when practice effects are present.

Results

As seen in Table 1, the patients and controls did not differ in age, sex, or education. However, healthy controls displayed significantly higher premorbid functioning as measured by a multiple choice vocabulary test (Lehrl, 1995). As expected, healthy participants exhibited significantly better memory performance than patients with schizophrenia. Patients' psychiatric and medication status remained stable over time—Brief Psychiatric Rating Scale total score: 41.0 (13.5), post:

38.2 ($SD = 14.9$), $t = 1.23$, $p > .2$; medication baseline: 358.3 mg (202.2), post: 358.3 (224.7), $t = .00$, $p > .9$.

RAVLT test-retest reliability coefficients for healthy controls and patients with schizophrenia were as follows: Trial 1 (healthy: $r = .39$; schizophrenia: $r = .41$), Trial 2 (healthy: $r = .51$; schizophrenia: $r = .49$), Trial 3 (healthy: $r = .57$, schizophrenia: $r = .63$), Trial 4 (healthy: $r = .53$, schizophrenia: $r = .73$), Trial 5 (healthy: $r = .51$, schizophrenia: $r = .58$), Interference Trial (healthy: $r = .32$; schizophrenia: $r = .34$), Trial 6 (healthy: $r = .73$, schizophrenia: $r = .63$), Trial 7 (healthy: $r = .63$; schizophrenia: $r = .41$), and Total Score (sum of Trials 1–5; healthy: $r = .72$; schizophrenia: $r = .69$). In both samples, no practice effects were evident—healthy participants, baseline: 59.2 (6.1) post: 57.8 ($SD = 8.5$; $t = 1.12$, $p > .2$); schizophrenia, baseline: 46.3 (12.3) post: 45.7 ($SD = 12.1$; $t = 0.44$, $p > .6$).

Reliable change estimates for the RAVLT total score were calculated for the schizophrenia sample. The SEM for time 1 was 6.83 and for time 2 was 6.71. The S_{diff} was 9.58. The 80%CI for measurement error was 12.26 words, and the 90%CI was 15.70 words. For the healthy participants, the SEM for time 1 was 3.22 and for time 2 was 4.50. The S_{diff} was 5.54. The 80%CI for measurement error was 7.09 words, and the 90%CI was 9.08 words.

The distribution of difference scores for the patients with schizophrenia was examined to determine if it conformed reasonably well to the RC calculations. Applying the 90%CI, based on the theoretical normal distribution, there should be 5% of participants falling in each tail. In the actual distribution, 2.6% of participants fell in each tail. Applying the 12-point cutoff score, which represents the .80CI, 7.9% of the sample improved and 13.2% declined.

Discussion

The goal of this study was to provide indexes of reliable change for memory performance in patients with schizophrenia to help clinicians determine if medication-related memory decline occurred. Baseline scores are best determined shortly before a patient is first taking a medication that might impair cognitive functioning (e.g., anti-parkinsonian agents). Retest scores should be assessed 2 weeks later. Baseline and retest assessment of an individual on the same medication, however, should be avoided, because detrimental effects to memory may already have occurred before baseline assessment.

None of the patients in this study received anti-parkinsonian agents or clozapine. This restriction was imposed to avoid distorting the range of probable change by including at-risk patients. Furthermore, our sample was medicated on both time points to provide scores that are unconfounded by changes in psychiatric status or medications. However, we suggest replication of our results with samples that are assessed while on wash-out to determine if results remain the same.

We used the sum of correctly recalled words from Trials 1 to 5 of the RAVLT as a learning and memory parameter, because composite scores are, in most cases, more reliable than single scores. In keeping with this principle, the test-retest reliability of the composite score in the sample of patients with schizophrenia ($r = .69$) and in the healthy controls ($r = .72$) was higher than most scores from single trials. Practice effects were absent in both the healthy controls and in the patients with schizophrenia. This is attributed to the fact that we used alternate forms of the RAVLT (versions A and B). Therefore, carryover effects resulting from identical material did not occur. In line with previous research (Aleman, Hijman, de Haan, & Kahn, 1999; Moritz, Heeren, Andresen, & Krausz, 2001), patients performed significantly worse than healthy controls on both occasions.

When using alternate forms of the RAVLT with patients with schizophrenia, a clinician can be 80% confident that an improvement or decline of 12 words in the total score is not the result of measurement error, and 90% confident that a change of 16 or more points in the total score is not the result of measurement error (for healthy participants, the corresponding values are 7 and 9). It is important to stress that the RC difference scores are meant to supplement, not replace, clinical judgment. It is clearly possible for patients to experience real decline or improvement even if their scores do not exceed the .80CI for measurement error. The clinician simply should have less confidence in clinical inferences based on changes that fall within the probable range of measurement error, and seek more ancillary evidence to support his or her opinion.

References

- Aleman, A., Hijman, R., de Haan, E. H. F., & Kahn, R. S. (1999). Memory impairment in schizophrenia: A meta-analysis. *American Journal of Psychiatry*, *156*, 1358–1366.
- Barr, W. B., & McCrea, M. (2001). Sensitivity and specificity of standardized neurocognitive testing immediately following sports concussion. *Journal of the International Neuropsychological Society*, *7*, 693–702.

- Buffett-Jerrott, S. E., & Stewart, S. H. (2002). Cognitive and sedative effects of benzodiazepine use. *Current Pharmaceutical Design*, 8(1), 45–58.
- Carter, J., Thrasher, S., & Thornicroft, G. (1994). Cognitive impairment and clozapine. *British Journal of Psychiatry*, 164(1), 132–133.
- Chelune, G. J., Naugle, R. I., Luders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, 7, 41–52.
- Grace, J., Bellus, S. B., Raulin, M. L., Herz, M. I., Priest, B. L., Brenner, V., et al. (1996). Long-term impact of clozapine and psychosocial treatment on psychiatric symptoms and cognitive functioning. *Psychiatric Services*, 47(1), 41–45.
- Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry*, 153(3), 321–330.
- Hageman, W. J. J. M., & Arrindell, W. A. (1993). A further refinement of the reliable change (RC) index by improving the pre–post difference score: Introducing RC(ID). *Behaviour Research and Therapy*, 31, 693–700.
- Hageman, W. J. J. M., & Arrindell, W. A. (1999a). Clinically significant and practical! Enhancing precision does make a difference. Reply to McGlinchey and Jacobson, Hsu, and Speer. *Behaviour Research and Therapy*, 37, 1219–1233.
- Hageman, W. J. J. M., & Arrindell, W. A. (1999b). Establishing clinically significant change: Increment of precision and the distinction between individual and group level of analysis. *Behaviour Research and Therapy*, 37, 1169–1193.
- Heaton, R. K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M. J., Marcotte, T. D., et al. (2001). Detecting change: A comparison of three neuropsychological methods using normal and clinical samples. *Archives of Clinical Neuropsychology*, 16, 75–91.
- Heubrock, D. (1992). Der Auditiv–Verbale Lerntest (AVLT) in der klinischen und experimentellen Neuropsychologie. Durchführung, Auswertung und Forschungsergebnisse [The Auditory–Verbal Learning Test (AVLT) in clinical and experimental neuropsychology. Administration, evaluation, and research findings]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 3, 161–174.
- Hinton-Bayre, A. D., Geffen, G. M., Geffen, L. B., McFarland, K. A., & Friis, P. (1999). Concussion in contact sports: Reliable change indexes of impairment and recovery. *Journal of Clinical and Experimental Neuropsychology*, 21, 70–86.
- Hsu, L. M. (1989). Reliable changes in psychotherapy: Taking into account regression toward the mean. *Behavioral Assessment*, 11, 459–467.
- Hsu, L. M. (1999). A comparison of three methods of identifying reliable and clinically significant client changes: Commentary on Hageman and Arrindell. *Behaviour Research and Therapy*, 37, 1195–1202.
- Iverson, G. (1999). Interpreting change on the WAIS–III/WMS–III following traumatic brain injury. *Journal of Cognitive Rehabilitation*, 17, 16–20.
- Iverson, G. I. (2001). Interpreting change on the WAIS–III/WMS–III in clinical samples. *Archives of Clinical Neuropsychology*, 16, 183–191.
- Iverson, G. L. (1998). Interpretation of Mini-Mental State Examination scores in community-dwelling elderly and geriatric neuropsychiatry patients. *International Journal of Geriatric Psychiatry*, 13, 661–666.
- Iverson, G. L., & Green, P. (2001). Measuring improvement or decline on the WAIS–R in inpatient psychiatry. *Psychological Reports*, 89, 457–462.
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336–352.
- Jacobson, & Revenstorf, D. (1988). Statistics for assessing the clinical significance of psychotherapy issues: Issues, problems, and new developments. *Behavioral Assessment*, 10, 133–145.
- Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology*, 67, 300–307.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Lee, M. A., Jayathilake, K., & Meltzer, H. Y. (1999). A comparison of the effect of clozapine with typical neuroleptics on cognitive function in neuroleptic-responsive schizophrenia. *Schizophrenia Research*, 37, 1–11.
- Lehrl, S. (1995). *Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B* (3rd ed.). Balingen: PERIMED-spitta.
- Lezak, M. D. (1995). *Neuropsychological assessment*. (3rd ed.). Oxford, England: Oxford University Press.
- Medalia, A., Gold, J., & Merriam, A. (1988). The effects of neuroleptics on neuropsychological test results of schizophrenics. *Archives of Clinical Neuropsychology*, 3, 249–271.
- Meltzer, H.-Y., & McGurk, S. (1999). The effects of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophrenia Bulletin*, 25, 233–255.
- Moritz, S. (2002). Cognitive dysfunctions. In D. Naber, M. Lambert, M. Krausz, C. Haasen, & D. Pickar, (Eds.), *Atypical neuroleptics in the treatment of schizophrenic patients* (pp. 167–175). Bremen, Germany: Uni-Med.
- 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(1), 23–31.
- Moritz, S., Krausz, M., Gottwalz, E., Lambert, M., Perro, C., Ganzer, S., et al. (2000). Cognitive dysfunction at baseline predicts symptomatic one-year outcome in first-episode schizophrenics. *Psychopathology*, 33, 48–51.
- Moritz, S., Perro, C., Woodward, T. S., & Krausz, M. (2002). Subjective cognitive dysfunction in first-episode patients predicts symptomatic outcome: a replication. *Psychopathology*, 35, 367–368.
- Ogles, B. M., Lambert, M. J., & Masters, K. S. (1996). *Assessing outcome in clinical practice*. Boston: Allyn & Bacon.
- Saint-Cyr, J. A., Taylor, A. E., & Lang, A. E. (1993). Neuropsychological and psychiatric side-effects in the treatment of Parkinson's disease. *Neurology*, 43(12, Suppl. 6), S47–S52.
- Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology*, 60, 402–408.
- Speer, D. C., & Greenbaum, P. E. (1995). Five methods for computing significant individual client change and improvement rates: Support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology*, 63, 1044–1048.

- Spohn, H. E., & Strauss, M. E. (1989). Relation of antipsychotic and anticholinergic medication to cognitive functions in schizophrenia. *Journal of Abnormal Psychology, 98*, 367–380.
- Sweeney, J. A., Keilp, J. G., Haas, G. L., & Weiden, P. J. (1991). Relationships between medication treatments and neuropsychological test performance in schizophrenia. *Psychiatry Research, 37*, 297–308.

- Temkin, N. R., Heaton, R. K., Grant, I., & Dikmen, S. S. (2000). Reliable change formula query: Temkin et al. reply. *Journal of the International Neuropsychological Society, 6*, 364.

Original submission June 24, 2002

Accepted December 3, 2002

Copyright of Applied Neuropsychology is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.