



The Word Memory Test and the One-in-Five-Test in an Analogue Study with Russian Speaking Participants

Stefan Tydecks.¹; Thomas Merten.²⁺; and Joe Gubbay.³

¹ *Humboldt University Berlin, Department of Psychology, Berlin, Germany;* ² *Klinikum im Friedrichshain, Department of Neurology, Berlin, Germany;* ³ *Private practice, Sydney, Australia*

Abstract

Symptom validity tests (SVTs) are increasingly employed throughout the world. This study tested the applicability of two SVTs with Russian-speaking participants using an analogue design. Russian versions of the Word Memory Test (WMT) and the One-in-Five-Test were given to 20 well educated young adults with an age range from 19 to 40 years (mean: 27.9). Three neuropsychological tests, the Trail Making Test (TMT), the Standard Progressive Matrices (SPM), and the Complex Figure Test (CFT), were also employed. Half of the sample received detailed instructions to feign memory deficits in order to obtain financial compensation within a litigation context, while the other half were to perform at their best. While a correct overall group classification of 100% was obtained with the WMT and 85% with the One-in-Five-Test, previously developed symptom validity indicators computed from TMT and SPM test scores did not perform well. While this appears to be the first study on modern SVTs undertaken in the Russian language, more work is needed from a cross-cultural perspective in the areas of test development and validation.

Keywords: Malingering; symptom validity testing; neuropsychological assessment; negative response bias

INTRODUCTION

While today there is little debate, in a number of English-speaking countries, about the usefulness and appropriateness of symptom validity testing in clinical-forensic neuropsychology, the same is not true for most European countries. This is reflected both by the small number of relevant publications from Europe and by the limited number of tests available for non-English language users. While some research activities are documented for certain languages such as Dutch (e.g., Schmand, Lindeboom, Schagen, Heijt, Koene, & Hamburger, 1998), German (e.g., Brockhaus & Merten, 2004),

and Spanish (e.g., Bilbao, Bize, & Bombín, 2002), the same is not true for a number of other languages. The need for access to symptom validity tests (SVTs) outside the English-speaking world is reflected by the number of foreign-language adaptations of the Word Memory Test (WMT; questions that can cue the respondent to the Green, 2003) which became available within a very short period of time. In December 2004, apart from the English original, there were also computerized WMT versions in Danish, Dutch, German, French, Spanish and Turkish, and oral versions in Hebrew, Italian, Mandarin, Portuguese and Russian.

⁺Address for correspondence: Richard A. Bryant, School of Psychology, University of New South Wales, NSW, 2052, Australia; Ph: +61-02-9385-3640; Facsimile: +61-02-9385-3541; Email: r.bryant@unsw.edu.au

The current study is the first to explore the validity of a Russian-language adaptation of the WMT, using an analogue design. Furthermore, a new SVT, the One-in-Five-Test, developed by the third author, was tested in a non-English language context for the first time.

Specifically, it was hypothesised that SVTs would be effective in discriminating between Russian-speaking subjects instructed to feign memory impairment and controls. Measures extracted from selected standard cognitive tests were also predicted to differentiate between groups.

METHOD

Participants

Twenty-one healthy, well-educated native speakers of Russian participated in the experiment. After the exclusion of one participant (see below), the final group consisted of 5 males and 15 females with a mean age of 27.9 years ($SD = 5.4$; range: 19 to 40 yr.) and a mean of 16.0 years of education ($SD = 2.7$). At the time of testing, all participants were staying in Germany, either for a limited professional stay or as long-term or permanent residents. They had resided in Germany for between three months and 10.5 years ($M = 5.3$ yr., $SD = 3.9$ yr.). One subject in the experimental group was excluded because he admitted in the post-experimental interview that he had not performed below his actual level although he had not had any difficulties understanding the malingering role outlined in the scenario.

Procedures

Participants were divided into two groups of 10 persons each on the basis of a pre-established randomized plan. The two groups did not differ significantly in age, education, or the time spent in Germany.

This study used an analogue design that was previously developed for testing German adaptations of various SVTs (Merten, Green, Henry, Blaskewitz, & Brockhaus, 2005; Merten, Henry, & Hilsabeck, 2004). With the exception of some of the instruments used in the study, the experimental design and scenario were an exact Russian-language adaptation of the procedure used by Merten et al. (2005) with German speaking participants.

The experimental malingerers received a detailed scenario designed to acquaint them with the role of a person who had suffered a car accident

two years before and was preparing for forensic neuropsychological assessment. After the accident, the person had lived through a period of relatively mild cognitive problems that had gradually improved with time. Moreover, the legal process to obtain compensation had been delayed for a long time. However, the claimant thought it was justified to receive damages for the handicaps and the social impediment suffered in the months following the accident. This, however, was only possible if symptoms could be demonstrated, in a believable way, during a neuropsychological assessment ordered by the court.

The scenario also contained a strong element of warning. The claimant was warned by his or her attorney not to overdo symptom presentation or else he or she might fail testing by an experienced neuropsychologist.

To ensure role understanding, a Russian translation of a 10-item multiple-choice questionnaire was given to the participants (see Appendix). At a minimum, 80 percent of the 10 items had to be answered correctly in order to proceed. In case of insufficient performance, the subject was asked to study the scenario carefully for a second time and to repeat the questionnaire. Apart from ensuring role commitment, the questionnaire was considered part of the coaching procedure.

In contrast, controls received a short instruction to put forth full effort in the testing session.

The experimental group was told that they would win a prize in the form of a book voucher of Euro 25 if they managed to fake neuropsychological impairment in a believable way. The controls were informed that the person with the best performance would be awarded the same book prize.

To check role commitment after the test session, a separate set of questions was answered by the experimental group. Finally, both the members of the experimental group and the controls were asked whether, in their opinion, some tests were too easy and whether there had been tests included which might have exclusively served to detect malingering or symptom exaggeration.

Testing time was approximately 70 to 90 minutes. The study was conducted from August to October 2004.

Instruments

The study was conducted by the first author in the Russian language. Tests were individually

administered according to standard directions for published tests in the following order:

1. The oral version of the WMT (Green, 2003), Immediate Recognition (IR) subtest.
2. The Standard Progressive Matrices (SPM; Raven, 1956) with a 20-minute time limit. Scoring of the SPM included a check of the performance curve (rate of decay) according to the procedure proposed by Gudjonsson and Shackleton (1986). The detailed cutoffs developed by McKinzey, Podd, Krehbiel, and Raven (1999) were used.
3. The copy trial of Rey's Complex Figure Test (CFT; Osterrieth, 1945).
4. The Trail Making Test (TMT; Reitan, 1992). A proposed malingering measure (ratio TMT-B : TMT-A) was computed using a cutoff of 1.49 / 1.50 (Iverson, Lange, Green, & Franzen, 2002).
5. The WMT Delayed Recognition (DR), Multiple Choice Recognition (MCR), Paired Associates (PA), and Free Recall (FR) subtests.
6. The oral version of the One-in-Five-Test. This test is adapted from the Digit Memory Test (Hiscock & Hiscock, 1989). Whereas SVTs typically result in scores around 50% correct if performed andomly, the One-in-Five-Test is designed so that a chance performance results in a score

around 80% correct. It is divided into 3 blocks of 12 trials. The subject is presented with a stimulus card containing 4 numbers for 5 seconds, and is instructed to remember the numbers. Following a delay, a response card containing all 4 stimulus numbers and one foil number is presented. The subject is required to identify any one number on the response card that was also on the stimulus card. Blocks 1 to 3 differ only in the delay following the stimulus card, which is 5, 10, and 15 seconds, respectively. Preliminary research by the third author found the One-in-Five-Test to be over twice as sensitive as the Digit Memory Test using a stringent below chance cutoff, among a sample of 30 cases referred for medicolegal neuropsychological or psychological assessment and judged to be exaggerating memory impairment. The total score across the 36 trials was used.

7. The 30-minute recall trial of the CFT.
8. The WMT, Long Delayed Free Recall (LDFR) subtest.

RESULTS

Results obtained for the two groups are presented in Tables 1 and 2. All cutoffs employed in this study are listed in Table 2. The test scores differed significantly for the two groups. The experimental malingerers scored lower on all tests.

Effect sizes (Z scores) for the SVTs and the performance measures were considerable. The exception was the TMT-B : TMT-A ratio, which did not significantly differentiate between experimental malingerers and controls.

Table 1
Raw Scores Obtained for two Groups on all Measures

| Instrument | Controls (<i>n</i> = 10) | | Experimental Malingers (<i>n</i> = 10) | | <u>U</u> Test (<i>Z</i>) |
|---|------------------------------|------|---|-------|-------------------------------|
| | <u>M</u> | SD | <u>M</u> | SD | |
| <u>Symptom Validity Tests</u> | | | | | |
| One-in-Five-Test | 36.0 | 0.0 | 27.4 | 9.1 | -3.73* |
| WMT—IR (%) | 97.0 | 2.3 | 63.3 | 24.0 | -3.78* |
| WMT—DR (%) | 98.8 | 1.8 | 52.8 | 29.0 | -3.65* |
| WMT—C (%) | 97.0 | 2.8 | 64.5 | 16.3 | -3.84* |
| <u>Symptom Validity Parameters Derived from Other Tests</u> | | | | | |
| SPM | non applicable | | | | |
| TMT-B : TMT-A ratio | 2.4 | 0.4 | 2.6 | 0.9 | -0.15 |
| <u>Neuropsychological Tests</u> | | | | | |
| WMT Multiple Choice (%) | 97.5 | 4.0 | 41.5 | 27.0 | -3.68* |
| WMT Paired Associates (%) | 95.5 | 9.5 | 44.0 | 21.0 | -3.75* |
| WMT Free Recall (%) | 85.8 | 14.5 | 26.8 | 16.8 | -3.68* |
| WMT Long Delayed (%) | 92.3 | 8.3 | 23.5 | 16.0 | -3.79* |
| SPM | 49.0 | 5.7 | 29.2 | 11.0 | -3.52* |
| CFT-IR | 27.7 | 2.6 | 19.5 | 6.5 | -3.18* |
| CFT—DR | 16.7 | 5.4 | 7.8 | 3.7 | -3.10* |
| TMT-A (seconds) | 32.4 | 5.7 | 99.4 | 58.9 | -3.79* |
| TMT-B (seconds) | 75.0 | 8.3 | 257.6 | 162.6 | -3.78* |

Note. WMT Word Memory Test, IR Immediate Recognition, DR Delayed Recognition, C Consistency; SPM Standard Progressive Matrices; TMT Trail Making Test; CFT Complex Figure Test, IR Immediate Recall, DR Delayed Recall.
* *p* < 0.05.

Table 2
Observed Variation and Quality of Classification for Symptom Validity Tests and Indicators

| Instrument | Cutoff | Controls | | | Experimental Malingers | | | Correct Classification (Total) |
|------------------|-------------|----------|------|--------|------------------------|------|--------|--------------------------------------|
| | | Range | Hits | Misses | Range | Hits | Misses | |
| One-in-Five-Test | 33 / 34 | 36 | 10 | 0 | 10-36 | 7 | 3 | 85% |
| WMT-IR | 82.5 / 85 | 92.5-100 | 10 | 0 | 7.5-92.5 | 9 | 1 | 95% |
| WMT-DR | 82.5 / 85 | 95-100 | 10 | 0 | 5-97.5 | 9 | 1 | 95% |
| WMT-C | 82.5 / 85 | 90-100 | 10 | 0 | 35-87.5 | 9 | 1 | 95% |
| WMT total | | | 10 | 0 | | 10 | 0 | 100% |
| SPM ¹ | | 42-59 | 7 | 1 | 11-45 | 4 | 5 | 65% |
| TMT-B:TMT-A | 1.49 / 1.50 | 1.7-2.9 | 10 | 0 | 1.4-4.0 | 1 | 9 | 55% |

Note. WMT Word Memory Test, IR Immediate Recognition, DR Delayed Recognition, C Consistency; TMT Trail Making Test; SPM Standard Progressive Matrices.

¹ Incomplete data sets for the rate-of-decay method because of incomplete test protocols in three participants, due to the 20-minute time limit.

Table 1

One subject in the experimental group scored higher than the proposed cutoff of 33 / 34 (82.5 / 85 %) on the WMT-IR trial, another on the WMT-DR, and a third on the WMT Consistency measure. Thus, for the WMT, the final group assignment was perfect.

The One-in-Five-Test showed a satisfying overall correct classification rate of 85%. Notably, no one in the control group had any difficulty with the task, all achieving perfect scores.

In contrast to these two tests, results for the TMT ratio and the SPM performance curve were poor. The total classification accuracy for both the SPM rate of decay and the TMT-B : TMT-A ratio was in the chance range (66% and 55%, respectively).

When analyzing participants' answers to the question about which instruments they thought might be too easy so their underperformance would become obvious, the results were as follows: No participant considered the WMT too easy. The One-in-Five-Test was named 10 times, the TMT twice and the CFT once. When examining answers to the question about which tests might constitute effort measures, the findings were somewhat different: The One-in-Five-Test was named 6 times as an effort measure, the WMT 4 times, and the CFT and the TMT once each.

With respect to below-chance performance, 3 experimental malingerers (30%) showed such a response pattern on at least one of the WMT recognition trials. For the One-in-Five-Test, 3 subjects scored below chance across the 36 trials. Importantly, only 2 malingerers were detected on both the WMT and One-in-Five-Test, with 2 subjects failing on only one test.

DISCUSSION

To the authors' knowledge, this is the first data obtained with a Russian-language adaptation of modern SVTs. The present analogue design produced results with the Russian WMT version similar to those which have been obtained with English, Turkish, and German language versions (Brockhaus & Merten, 2004; Brockhaus & Peker, 2003; Iverson, Green, & Gervais, 1999; Merten et al., 2004; Tan, Slick, Strauss, & Hultsch, 2002). As in the present study, all these studies report excellent classification results for the symptom validity parameters of the WMT. This suggests that the Russian version shares the test characteristics of the original WMT and other foreign language adaptations. Further research using larger samples, however, is required.

Research is required to demonstrate the validity of a test that is adapted for a population differing in some characteristic from the original target population (e.g., Poortinga, 1995). There is, however, no evidence to suggest that symptom validity testing does not operate in a Russian-speaking population in the same way it does with other populations. The hypothesis that SVTs would be sensitive to malingering among Russian-speakers is supported by the present results. The study was based on a small sample, but large effect sizes similar to those obtained in previous studies (Merten et al., 2004, 2005) were expected. It has to be borne in mind, however, that reliable estimates of sensitivity and specificity cannot be based on such low sample sizes.

The One-in-Five-Test is a new SVT. As is the case with a number of digit memory SVTs (e.g., Hiscock & Hiscock, 1999; Allen, Conder, Green, & Cox, 1997), the task is easy to perform, with low scores suggestive of insufficient effort or frank malingering. Importantly, all controls obtained the maximum score in the test. It remains to be demonstrated that the task is also simple for individuals with genuine cognitive impairment, although the third author's experience is that scores below 92% should be treated with caution. Even though the sample was composed of younger, well educated individuals, 40% of the experimental group performed below chance on at least one SVT. As several authors (e.g., Cato, Brewster, Ryan, & Giuliano, 2002) have pointed out, it is advisable to administer more than one SVT as feigning of impairment is often not a consistent behavior, a suggestion that appears to be borne out in the current study. In this vein, Iverson (2003) wrote that

"Many clinicians now include one or more specialized tests in their standard battery. Some clinicians include numerous specialized tests, and examine the standard test results carefully for suspicious performance patterns." (p. 168)

In contrast, two purported malingering indicators derived from the TMT and SPM did not perform well. Similar results have been found by Merten et al. (2004) and Merten et al. (2005) with German-speaking populations. Although it is true that simulation studies are not sufficient to validate SVTs, given the nature of the behavior being assessed, analogue studies will remain an important component in the validation process, along with clinical validity studies.

It is likely that symptom validity testing will become increasingly important for Russian psychology in the next few years. The availability of Russian test versions is also of great interest for

neuropsychologists in Europe and Northern America. Since the early 1970s, over 400,000 Russian-speaking persons have immigrated to the United States (Birman & Trickett, 2001). Germany is another country with a high incidence of immigrants from the former Soviet Union; by 2003, a sum total of 343,391 immigrants with Russian as the mother tongue (e.g., Russian Federation, Belarus) was reported (Statistisches Bundesamt, 2004).

In conclusion, the current study provides preliminary evidence supporting the validity of Russian adaptations of SVTs including the WMT and One-in-Five-Test in detecting feigning of memory impairment.

REFERENCES

- Allen, L. M., Conder, R. L., Green, P., & Cox, D. R. (1997). *CARB '97 Manual for the Computerized Assessment of Response Bias*. Durham, NC: CogniSyst.
- Bilbao, A., Bize, A., & Bombin, I. (2002). Factors influencing the performance of simulators of memory deficits [Abstract]. *Journal of the International Neuropsychological Society*, 8, 485.
- Birman, D., & Trickett, E. J. (2001). Cultural transitions in first-generation immigrants: Acculturation of Soviet Jewish refugee adolescents and parents. *Journal of Cross-Cultural Psychology*, 32, 456-477.
- Brockhaus, R., & Merten, T. (2004). Neuropsychologische Diagnostik suboptimalen Leistungsverhaltens mit dem Word Memory Test [Neuropsychological assessment of suboptimal performance: the Word Memory Test]. *Nervenarzt*, 75, 882-887.
- Brockhaus, R., & Peker, Ö. (2003). Testing Effort in Turkish-speaking Subjects: Validation of a Translation of the Word Memory Test (WMT) [Abstract]. *Journal of the International Neuropsychological Society*, 9, 520.
- Cato, M. A., Brewster, J., Ryan, T., & Giuliano, A. J. (2002). Coaching and the ability to simulate mild traumatic brain injury symptoms. *The Clinical Neuropsychologist*, 16, 524-535.
- Green, P. (2003). *Green's Word Memory Test for windows: User's manual*. Edmonton, Canada: Green's Publishing.
- Gudjonsson, G. S., & Shackleton, H. (1986). The pattern of scores on Raven's Matrices during 'faking bad' and 'non-faking' performance. *British Journal of Clinical Psychology*, 25, 35-41.
- Hiscock, M., & Hiscock, D. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology*, 11, 967-974.
- Iverson, G. L. (2003). Detecting malingering in civil forensic evaluations. In A. M. Horton, L. C. Hartlage (eds.), *Handbook of Forensic Neuropsychology* (137-177). New York: Springer.
- Iverson, G. L., Green, P., & Gervais, R. (1999). Using the Word Memory Test to detect biased responding in head injury litigation. *Journal of Cognitive Rehabilitation*, 17, 4-8.
- Iverson, G. L., Lange, R. T., Green, P., & Franzen, M. D. (2002). Detecting exaggeration and malingering with the Trail Making Test. *The Clinical Neuropsychologist*, 16, 398-406.
- McKinzev, R. K., Podd, M. H., Krehbiel, M. A., & Raven, J. (1999). Detection of malingering on Raven's Standard Progressive Matrices: A cross-validation. *British Journal of Clinical Psychology*, 38, 435-439.
- Merten, T., Green, P., Henry, M., Blaskewitz, N., & Brockhaus, R. (2005). Analog validation of German language symptom validity tests and the influence of coaching. *Archives of Clinical Neuropsychology*, 20, 719-726.
- Merten, T., Henry, M., & Hilsabeck, R. (2004). Symptomvalidierungstests in der neuropsychologischen Diagnostik: eine Analogstudie [The use of symptom validity tests in neuropsychological assessment: an analogue study]. *Zeitschrift fuer Neuropsychologie*, 15, 81-90.
- Osterrieth, P.-A. (1945). Le test de copie d'une figure complexe. Contribution à l'étude de la perception et de la mémoire [The Complex Figure Copying Test. Contributions to the study of perception and memory]. *Archives de Psychologie*, 30, 205-353.
- Poortinga, Y. H. (1995). Cultural bias in assessment: Historical and thematic issues. *European Journal of Psychological Assessment*, 11, 140-146.
- Raven, J. C. (1956). *Standard Progressive Matrices, Sets A, B, C, D, E*. London: Lewis.
- Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring*. South Tucson, AZ: Reitan Neuropsychology Laboratory.
- Schmand, B., Lindeboom, J., Schagen, S., Heijt, R., Koene, T., & Hamburger, H. L. (1998). Cognitive complaints in patients after whiplash injury: the impact of malingering. *Journal of Neurology, Neurosurgery, and Psychiatry*, 64, 339-343.

- Statistisches Bundesamt (2004). *Bevölkerung: Ausländische Bevölkerung nach Staatsangehörigkeit 2003* [Population: Foreign population in 2003, split by nationalities]. Wiesbaden, Germany: Author.
- Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist, 16*, 495-505.

APPENDIX

Questionnaire Concerning the Scenario (Comprehension of Role to be Assumed)
(Only one answer is correct)

1. What kind of problems did you experience after the accident?
 - a) paralysis
 - b) memory problems
 - c) epileptic fits
2. Were you able to remember a newspaper article after having read it?
 - a) yes, always
 - b) sometimes
 - c) no, never
3. With what kind of memory function did you have no major problems?
 - a) important personal facts from the past (before the accident took place)
 - b) shopping lists
 - c) important events after the accident
4. Did the accident affect your professional and private life?
 - a) Yes, I received more support from colleagues and friends after the accident than I had previously.
 - b) Yes, private and professional relationships suffered.
 - c) No, there were no major changes.
5. What was your ability to concentrate like after the accident?
 - a) completely normal, in contrast to my memory problems
 - b) I couldn't concentrate on even the simplest things such as reading a newspaper article any more.
 - c) I was not able to concentrate on things as well as I used to.
6. What is your goal during this upcoming evaluation?
 - a) I want to show maximum difficulties with memory and concentration so that they will believe me.
 - b) I want to take pains to present my actual symptoms after the accident.
 - c) I want to be as good as possible in the tests so that I will appear credible.
7. What could possibly happen if you exaggerate the symptoms to a degree greater than you originally experienced them?
 - a) The examiner would not believe me and I would not receive compensation.
 - b) The compensation would be greater than I feel that I deserve.
 - c) There would be no consequences.
8. If you wanted to be credible during this examination, you would
 - a) say that you can't remember anything at all any more.
 - b) maintain that you cannot remember having had any memory problems at all.
 - c) describe that the gravity of the memory problems has been fluctuating.
9. Do you think that a qualified examiner will be able to distinguish between real and exaggerated memory deficits?
 - a) No, I don't think so.
 - b) Perhaps, but it is very unlikely that I will get caught.
 - c) Yes, I am going to have to be careful.

10. Do you feel that it is fair to want to convince the evaluator that you have memory problems?
- a) Yes, because I am still suffering from serious problems with memory and concentration.
 - b) Yes, because I really did suffer from these symptoms and it is not my fault that the court case was dragged out for such a long time.
 - c) No, and for this reason I am going to take the exam to the best of my ability.