Effects of Korsakoff Amnesia on performance and symptom validity testing

Erik Oudman, Emmy Krooshof, Roos van Oort, Beth Lloyd, Jan W. Wijnia & Albert Postma

Routledge
Taylor & Francis Group

Check for updates

Effects of Korsakoff Amnesia on performance and symptom validity testing

Erik Oudman[a,b] (iD), Emmy Krooshof[a,b], Roos van Oort[c], Beth Lloyd[a,b], Jan W. Wijnia[a,b], and Albert Postma[a,b]

[a]Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands; [b]Slingedael Korsakoff Center, Lelie Care Group, Rotterdam, The Netherlands; [c]Wettstein & Peterse Expertise (WPEX), Amersfoort, The Netherlands

ABSTRACT

Performance validity tests (PVTs) and Symptom validity tests (SVTs) are developed to identify people that present false or exaggerated symptoms. Although a key factor of both types of tests includes relative insensitivity to cognitive disorders, the direct effects of amnesia have been poorly studied. Therefore, a sample of 20 patients diagnosed with Korsakoff Amnesia (KA) through neuropsychological assessment and 20 healthy comparisons (HC) were administered the Test of Memory Malingering (TOMM), the Structured Inventory of Malingered Symptomatology (SIMS), and the newly developed Visual Association Test – Extended (VAT-E). Our results show that KA patients scored systematically lower on the TOMM and VAT-E, while performance on the SIMS was comparable with healthy comparisons. Some KA patients were regarded as underperformers based on the TOMM and VAT-E, suggesting limitations in applying these instruments in severe amnesia. There was a strong interdependence of PVTs in logistic regression. We conclude that the TOMM and VAT-E are not fully robust against severe memory disorders and show a serious risk of false positives. Complete neuropsychological profile analysis is needed, and PVTs should be interpreted with caution in patients with suspected amnesia.

Introduction

When using cognitive tests for diagnostic purposes it is crucial to know whether the patient is exaggerating or fabricating cognitive deficits. Between 22 and 40% of all the patients involved in litigation have noncredible performance on neuropsychological assessments (Mittenberg, Patton, Canyock, & Condit, 2002; Rogers, 2008). Performance validity tests (PVTs) are tests for assessing the credibility of the patient's performance on neuropsychological assessments, and symptom validity tests (SVTs) are questionnaires to index exaggeration of symptoms (Bush et al., 2005). Failure on a PVT or SVT reflects contributions to test performance other than neurological or cognitive factors (Bigler, 2012). This is particularly the case when memory symptoms are feigned or effort was too little.

Recently neuropsychologists have become increasingly aware that, although cognitive testing is highly sensitive to brain dysfunction, poor test results are nonspecific and can arise for a variety of reasons (Greher & Wodushek, 2017). Not surprisingly, the use of PVTs and SVTs has climbed dramatically in clinical practice (Sweet, Benson, Nelson, & Moberg, 2015).

PVTs are standalone or embedded tests designed and administered for the purpose of determining invalid responding. At face value, they seem to be similar to other neuropsychological tests, although they are largely insensitive to brain impairments. Practically speaking, cognitively impaired patients are able to pass these tests despite the severity of their cognitive problems (Wodushek & Greher, 2017). Some PVTs have fixed cutoff scores, but others are based on hierarchical profile analysis of the PVT in relation to other neuropsychological tests. For example, in the Medical Symptom Validity Test, performance on "easy" and "hard" subtests is directly compared, to reduce false positives (Wodushek & Greher, 2017). SVTs are self-report questionnaires that are designed to detect feigned or exaggerated psychopathology. The general idea is that respondents are likely to endorse bizarre, rare, atypical, or extreme symptoms on a questionnaire when they attempt to feign or exaggerate symptoms (Van Impelen, Merckelbach, Jelicic, & Merten, 2014).

Many PVTs and SVTs have been developed and validated over the years. The most frequently applied,

and probably the most validated PVT is the TOMM (Dandachi-FitzGerald, Ponds, & Merten, 2013; Martin, Schroeder, & Odland, 2015). The TOMM is based on visual stimuli where the patient has to remember and recognize 45 out of 50 individual pictures (Tombaugh, 1996). Since the TOMM has become a very popular PVT, with even online tutorials how to pass this test as a malingerer (Wodushek & Greher, 2017), there is an urgent need to develop new PVTs. Moreover, in new PVTs it would be relevant to combine assessment of declarative memory, and validity testing in one test. Recently, the Visual Association Test-Extended (VAT-E) was developed as a new PVT that is also able to detect declarative memory problems in subtests. The VAT-E forms an extension of the traditionally used Visual Association Test (VAT) which is an episodic memory test (Lindeboom, Schmand, & Christensen, 2003; Meyer, de Jonghe, Schmand, & Ponds, 2017), and yields very promising results in discriminating patients with Alzheimer's disease, Mild Cognitive Impairment (MCI), and healthy persons that feign memory deficits (Meyer et al., 2017). In the VAT-E both fixed cutoff scores are combined with hierarchical profile analysis, comparing scores on multiple indices. A specifically prominent SVT that is developed is the Structured Inventory of Malingered Symptomatology (SIMS) (Smith & Burger, 1997). In contrast to PVTs, the SIMS is not developed to indicate underperformance on neuropsychological assessments, but over-reporting of symptomatology. Dandachi-FitzGerald, Ponds, Peters, and Merckelbach (2011) emphasized on the importance to test both aspects of symptom exaggeration, since over-representation of symptoms was not necessarily related to underperformance. A thorough assessment of validity therefore needs to take both dimensions into account.

An important question is whether PVTs and SVTs are able to discriminate patients with actual cognitive deficits from patients that feign symptoms. Currently, the literature is inconclusive on this matter, with some studies suggesting that the majority of PVTs are not always suitable for patients with severe cognitive deficits (Bortnik, Horner, & Bachman, 2013; Dorociak, Schulze, Piper, Molokie, & Janecek, 2018; Rudman, Oyebode, Jones, & Bentham, 2011; Walter, Morris, Swier-Vosnos, & Pliskin, 2014, see McGuire, Crawford, & Evans, 2019 for a review), whereas other studies find that PVTs can be safely used in a population with severe cognitive deficits (Hampson, Kemp, Coughlan, Moulin, & Bhakta, 2014; Liu et al., 2016; Meyer et al., 2017; Slick et al., 2003; Tombaugh, 1996, 1997). A point of criticism regarding the influence of

cognitive disorders on PVT performance is that cognitive functioning is directly related to PVT performance in some studies on dementia. Rudman et al. (2011) found that specifically learning of new information was significantly correlated with task performance on the frequently applied Test of Memory Malingering (TOMM), and the Medical Symptom Validity Test. Selective attention was also significantly related to the Coin in hand test. The authors argued that the specificity of the applied PVTs was not optimal, since some moderate to severe dementia patients were classified as displaying suboptimal effort. Bortnik et al. (2013) made even stronger claims, stating that the majority of effort tests demonstrated unacceptably high false-positive rates in their sample of 164 dementia patients. In fact, 22% of the patients were unable to pass the second trial of the TOMM, being misclassified as underperformers. In contrast, some normative data suggest that the effects of cognitive disorders are rather limited (Tombaugh, 1997).

Although PVTs are widely used in a variety of settings to differentiate feigned and actual memory issues, there are currently very limited studies available on the validity of PVTs in patients with profound memory disorders with relatively preserved other functions. Earlier investigations on PVT performance in amnesia frequently incorporated patients in which memory was not the only or central cognitive problem, but part of general cognitive dysfunction (Loring, Larrabee, Lee, & Meador, 2007; Tombaugh, 1997). In the present study, we enrolled patients with Korsakoff Amnesia (KA). KA is a chronic brain disorder, caused by thiamin (vitamin B1) deficiency. In the industrialized world, the most common cause of thiamin deficiency is severe alcoholism. The most essential symptom of KA is profound declarative amnesia for learning and remembering new material (Kopelman, Thomson, Guerrini, & Marshall, 2009). Commonly, but not necessarily, executive deficits are present, such as problems with inhibition of behavior, high interference of information sensitivity, poor judgment, poor planning abilities, problem solving inabilities, and perseverative responses (Moerman - van den Brink et al., 2019). In contrast to many forms of dementia, apraxia, aphasia, agnosia, and broad intellectual decline are not symptoms of KA (Kopelman et al., 2009), increasing the relevance of testing the sensitivity of PVTs and SVTs in KA.

The aim of this study was to test whether patients with severe amnesia are able to successfully pass the TOMM, SIMS, and VAT-E. Moreover, we wanted to elucidate how the outcome of the TOMM, SIMS, and

Table 1. Demographic characteristics.

	Healthy Comparisons ($n = 20$)	Korsakoff Amnesia ($n = 20$)	
Number of females	12	10	χ^2 (1, $n = 40$) = 0.40, $p = .53$
Age[a] (M, SD)	55.4 (4.9)	57.8 (7.3)	$t(38) = 1.14$, $p = .260$, $\eta^2 = .04$
Premorbid IQ scores (M, SD)[b]	97.7 (10.9)	93.6 (11.4)	$t(38) = 1.22$, $p = .230$, $\eta^2 = .04$
VAT-E Paired Association (M, SD)[c]	16.6 (5.6)	5.6 (4.3)	$t(38) = 6.77$, $p < .001$, $\eta^2 = .55$
VAT-E Free Recall (M, SD)[c]	23.8 (6.6)	9.0 (5.2)	$t(38) = 1.14$, $p < .001$, $\eta^2 = .62$
CAMDEX Nonmemory impaired %[d]		21.4	
CAMDEX Memory impaired %[d]		100.0	
MMSE impaired %[e]		73.3	
Digit Test Forward impaired %[f]		9.0	
Digit Test Backward impaired %[f]		9.0	
BADS impaired %[g]		25.0	

[a]Age in years.
[b]Premorbid IQ scores were estimated with the Dutch version of the National Adult Reading Test (Schmand et al.,1991).
[c]VAT-E subtest Paired Association and Free Recall are the memory measures of the VAT-E (Meyer & de Jonghe, 2017). Higher scores represent better performance on the subtest.
[d]CAMDEX total scores for the nonmemory and memory section, as an index of general cognitive functioning, normed according to Roth et al. (1986).
[e]MMSE total scores, as an index of general cognitive functioning, normed according to Folstein et al. (1975).
[f]Digit span, as a test for concentration and working memory, normed according to Lindeboom and Matto (1994).
[g]Behavioral Assessment of the Dysexecutive Syndrome (BADS), a test battery for executive functioning, normed according to Wilson et al. (1996).

VAT-E could predict the presence of KA, and whether combining evidence would reduce the effects of severe amnesia.

Materials and methods

Participants

Twenty patients diagnosed with Korsakoff Amnesia (KA) through extensive multidisciplinary observation and diagnostics participated in this study (see Table 1 for an overview). They were all inpatients of the Korsakoff Center "Slingedael," Lelie Care Group, in Rotterdam, The Netherlands. All patients fulfilled the DSM-5 criteria for the Alcohol-induced major neurocognitive disorder, Amnestic Confabulatory type (code: 291.1) (American Psychiatric Association, 2013), and the characteristics of KA described by Kopelman (2002). All patients had been sober for more than six months and required intensive sheltered living due to the severity of their amnesia. All patients fulfilled the D criterion of Slick et al. (2003) regarding the severity of genuine cognitive pathology and warranted intensive sheltered living based on the severity of their amnesia. The amnestic syndrome was confirmed by extensive neuropsychological testing. All patients were in the chronic, amnestic stage of the syndrome, none of the patients were in the confusional Wernicke psychosis at the moment of testing. Other lifetime exclusion criteria were illiteracy, presence of additional neurological disorders (traumatic brain injury, epilepsy, stroke, or brain tumor), acute psychiatric conditions (psychosis, major depression, etc.), and physical conditions interfering with the testing procedure. Twenty healthy comparisons (HC), statistically equivalent in age, IQ, and gender were also enrolled in the study. IQ was estimated with the Dutch version of the National Adult Reading Test (Schmand, Bakker, Saan, & Louman, 1991). The project was conducted according to the declaration of Helsinki and written informed consent was obtained for all participants.

Tasks and procedure

Test of Memory Malingering (TOMM)

In the TOMM, 50 pictures were shown sequentially to the participant. After the learning trial, the participant was presented with 50 forced choices between two pictures of which one of the two had been shown in the learning trial. After the first trial, the participant was shown the same pictures in a second learning trial. Lastly, the participant was asked once more to select the picture he/she had seen during the learning trial in 50 consecutive forced choices. During all trials the participant received feedback on their answer, as indicated by the TOMM manual. TOMM trial 2 was analyzed.

Visual Association Test-Extended (VAT-E)

For the VAT-E, the participant was instructed to look at and remember 24 pictures with one object, animal, or person on it, one by one for three seconds each, in two learning trials. In the next section the 24 pictures were shown again, this time with an additional item in the picture. The participant was instructed to name both items and identify which of the pictures had to be remembered during the learning trials (immediate recall; IR) to ensure the association of the pictures. After a 15-minute interval in which no memory or visual tests were administered, the participant was again shown the 24 pictures with the additional

picture and was again asked to identify which picture had to be remembered during the learning trial (delayed recall: DR). This was followed by the paired association (PA) in which the participant was shown the initial 24 pictures and had to name which picture was added. Next, the participant was asked to freely recall which of the pictures he/she still remembered (free recall; FR). These could be named in pairs or separately (e.g., either a candle on a football or a candle and a football). The last trial consisted of 12 of the 24 pictures shown individually again as a cue and the participant having to choose between four pictures which was the associated picture and was interpreted as the profile analysis score (FR $\geq$ 7 in combination with MC $\leq$ 9). The consistency subtest (CN) was derived from the IR and DR scores, and represented whether the participant had the same answers correct or wrong on the IR compared to the DR test. The subtests that measured underperformance were IR, DR, and CN, and the profile analysis score.

Structured Inventory of Malingered Symptomatology (SIMS)

The SIMS consisted of 75 items regarding symptomatology. If the amount of affirmative answers exceeded the cutoff value, it is an indication of feigning psychiatric symptoms (Smith & Burger, 1997). A score above 16 is indicative of aggravating.

Neuropsychological assessment

Premorbid IQ scores were estimated with the Dutch version of the National Adult Reading Test for all participants (Schmand et al., 1991). The VAT-E subtest Paired Association and Free Recall are the memory measures of the VAT-E (Meyer & de Jonghe, 2017) and were collected for all participants. Higher scores represented better performance on the subtest.

For the KA patients, the percentage of patients that scored below the cutoff value was reported for six additional indices of cognitive functioning: as an index for global cognitive functioning, the CAMDEX total scores (cutoff value < 79 points), the nonmemory CAMDEX score (cutoff < 62 points), and memory CAMDEX score (cutoff < 27 points) were reported (Roth et al., 1986; Schmand, Walstra, Lindeboom, Teunisse, & Jonker, 2000). Moreover, the MMSE total scores (cut of <24) was applied as an index of general cognitive functioning (Folstein, Folstein, & McHugh, 1975). As a test for concentration and working memory, the digit span was applied (see Lindeboom & Matto, 1994 for differential norms). Finally, the Behavioural Assessment of the Dysexecutive Syndrome (BADS), a test battery for executive functioning, was scored. A test score was considered "impaired" if the standard score was more than 1.5 SD below the normative mean) (Wilson, Alderman, Burguess, Emslie, & Evans, 1996).

Statistics

The TOMM, VAT-E, and SIMS were presented in a counterbalanced order. Because the assumptions for parametric testing were violated by skewness of the data, nonparametric U-tests were performed on the TOMM trial 1 and trial 2, VAT-E IR, DR, and CN, and SIMS, to index possible differences in median values between KA patients and controls. Bonferroni corrected p-values were presented. Moreover, Logistic regression was carried out as a hierarchical approach to investigate accumulative prediction of severe amnesia. Receiver Operated Characteristic curves were estimated and objectified as the Area Under the Curve (AUC) to index the ability of the tests to discriminate KA patients and healthy controls correctly. Also, Chi-squares were calculated for the proportions of controls and patients that were unable to pass the aforementioned tests, as well as the profile analysis score of the VAT-E, to indicate the effects of combining evidence.

Results

Table 1 shows a summary of neuropsychological tests, demographic variables, and the performance on background variables for both patients diagnosed with KA and the healthy comparisons. Both groups were statistically equivalent on age, gender, and IQ. As expected, KA patients scored significantly lower on the VAT-E memory indices, representing the severity of the amnesia in KA (see Table 1).

Prevalence of underperformance and over-reporting in Korsakoff Amnesia and healthy comparisons

Table 2 shows the median and range scores for the patients diagnosed with KA and the healthy comparisons. Task performance was significantly worse in KA patients for all trials of the TOMM and VAT-E, compared to healthy comparisons. The scores were comparable on the SIMS. Task performance on the PVTs, but not the SVT, were sensitive for KA.

Table 2. Group differences across Performance and Symptom Validity Tests in 20 patients diagnosed with Korsakoff Amnesia and 20 Healthy comparisons.

	Healthy Comparisons ($n = 20$)	Korsakoff Amnesia ($n = 20$)	Statistics	Bonferroni Corrected p-values
TOMM - Trial 1 (Median, Range)[a]	48.5 (41–50)	44.5 (24–49)	$U = 79.50$	$p < .001$
TOMM - Trial 2 (Median, Range)[a]	50.0 (49–50)	49.0 (42–50)	$U = 76.50$	$p < .001$
SIMS (Median, Range)[b]	4.0 (2–18)	6.0 (1–15)	$U = 155.50$	$p = 1.000$
VAT-E Immediate Recognition (Median, Range)[c]	24.0 (23–24)	24.0 (21–24)	$U = 118.50$	$p = .002$
VAT-E Delayed Recognition (Median, Range)[c]	24.0 (24)	22.5 (18–24)	$U = 70,00$	$p < .001$
VAT-E Consistency (Median, Range)[c]	24.0 (23–24)	23.0 (18–24)	$U = 75.50$	$p < .001$

Note. VAT-E = Visual Association Test – Extended (Meyer et al., 2017); TOMM = Test of Memory Malingering (Tombaugh, 1996); SIMS = Structured Inventory of Malingered Symptomatology (Smith & Burger, 1997).
[a]Median Test score on the first and second trial of the Test of Memory Malingering (Tombaugh, 1996). Higher scores represent better performance on the subtest.
[b]Test score on the Structured Inventory of Malingered Symptomatology (Smith & Burger, 1997). Higher scores represent more malingered symptomatology.
[c]VAT-E subtest Immediate Recognition, Delayed Recognition, and Consistency are the underperformance measures of the VAT-E (Meyer & de Jonghe, 2017). Higher scores represent better performance on the subtest.

Table 3. The number of Korsakoff amnesia patients ($n = 20$) and healthy comparisons ($n = 20$) failing on the Test of Memory Malingering (Trial 1 and 2), the Structured Inventory of Malingered Symptomatology (SIMS), and the subtests of the Visual Association Test – Expanded.

Single tests	Healthy Comparisons	Korsakoff Amnesia	
TOMM - Trial 1	3 (15%)	9 (45%)	$\chi^2(1) = 5.58, p = .02, \eta = .37$
TOMM - Trial 2	0 (0%)	2 (10%)	$\chi^2(1) = 2.11, p = .15, \eta = .23$
VAT-E Immediate Recognition	0 (0%)	1 (5%)	$\chi^2(1) = 1.02, p = .31, \eta = .16$
VAT-E Delayed Recognition	0 (0%)	5 (25%)	$\chi^2(1) = 5.71, p = .02, \eta = .38$
VAT-E Consistency Index	0 (0%)	5 (25%)	$\chi^2(1) = 5.71, p = .02, \eta = .38$
SIMS	1 (5%)	0 (0%)	$\chi^2(1) = 1.02, p = .31, \eta = .16$
Failing on both of the combined tests			
VAT-E Profile analysis	0 (0%)	0 (0%)	
TOMM – Trial 2 & Sims	0 (0%)	0 (0%)	
VAT-E – Immediate Recognition & SIMS	0 (0%)	0 (0%)	
TOMM – Trial 2 & VAT-E Immediate Recognition	0 (0%)	1 (5%)	$\chi^2(1) = 1.02, p = .31, \eta = .16$
TOMM – Trial 2 & VAT-E Delayed Recognition	0 (0%)	2 (10%)	$\chi^2(1) = 2.11, p = .15, \eta = .23$
TOMM – Trial 2 & VAT-E Consistency Index	0 (0%)	2 (10%)	$\chi^2(1) = 2.11, p = .15, \eta = .23$

Note. IQ = Intelligence Quotient based on the Dutch version of the National Adult Reading Test (Schmand et al., 1991); VAT-E = Visual Association Test – Extended (Meyer et al., 2017); TOMM = Test of Memory Malingering (Tombaugh, 1996); SIMS = Structured Inventory of Malingered Symptomatology (Smith & Burger, 1997); IR = Immediate Recognition; DR = Delayed Recognition; N.A. = Not Available because of ceiling scores on one of the indices.

Logistic regression

Linear regression collinearity diagnostics indicated strong multicollinearity of the data structure for the indices of the TOMM, VAT-E, and SIMS, with a condition index of >30 for four of the subscales. In a forward conditional logistic regression analysis with all variables, the best model fit was reached with only the VAT-E Delayed Recognition scale ($\chi^2(5) = 24.55$, $p < .001$), suggesting good model fit, with Cox and Snell R2 = .46, and a correct classification of 82.5%. This minimal model suggests strong interdependence of the PVT indices.

Receiver Operating Characteristic (ROC)-analysis

In ROC-analysis, the AUC of the SIMS is .61, suggesting a low ability to predict KA. The AUC was .80 for Trial 1 of the TOMM, and .81 for Trial 2 of the TOMM, suggesting that both subtests of the TOMM were good in predicting KA. The AUC was .70 for the immediate recognition, .83 for the delayed recognition, and .81 for the consistency score of the VAT-E,

suggesting reasonable to good ability to detect KA. These results suggest that both the TOMM and VAT-E subscales intended to index underperformance are good in predicting whether participants were in the amnesia patient or healthy comparison group.

Proportional analysis

Table 3 shows the proportions of patients and healthy comparisons who failed the TOMM, the VAT-E, and the SIMS, and who failed on two combined tests. As can be seen, 1–9 (5–45%) of the KA patients failed the subtests on one or both PVTs. The proportion of the two groups that failed on the TOMM – Trial 1 ($\chi^2(1) = 5.58$, $p = .02$, $\eta = .37$), the VAT-E Delayed Recognition ($\chi^2(1) = 5.71$, $p = .02$, $\eta = .38$), and VAT-E Consistency scale ($\chi^2(1) = 5.71$, $p = .02$, $\eta = .38$), were statistically different between KA patients and controls. The results represented in Table 3 suggest that for both the TOMM and VAT-E, the number of false positives were higher in the KA

patients than in healthy comparisons. Combining PVTs was not effective in eliminating false positives in KA patients, and had comparable false response rates as the best out of two elements. Importantly, performing profile analysis, or combining a PVT index with the SIMS was effective in eliminating false positives.

Discussion

The aim of our study was to identify whether patients with severe Korsakoff amnesia (KA) are able to pass Performance Validity Tests (PVTs) and Symptom Validity Tests (SVTs) despite the severity of their amnesia. Moreover, we wanted to elucidate how the outcome of the TOMM, SIMS, and VAT-E could predict the presence of KA, and whether combining evidence would reduce the effects of severe amnesia. Our results show that PVT task performance was lower for the TOMM and VAT-E in KA patients, and did lead to 5–45% false positive scores on the PVT subscales. In logistic regression, Z-transformed indices could predict KA with an accuracy of 82.5%, and the individual PVT indices all had reasonable to good ability to predict KA in ROC-analysis. Combining PVTs was not effective in reducing the number of false positives to zero, although profile analysis and including the SIMS as additional index were effective. Our results suggest that the TOMM and VAT-E should be interpreted with serious caution in patients with suspected severe amnesia, since the severity of amnesia does affect PVT performance.

The fundamental assumption of a PVT is insensitivity to cognitive dysfunction (Tombaugh, 1997). The results of the present study show that PVT performance is not insensitive to severe amnesia due to KA. Earlier studies in dementia highlighted specificity rates of 24% of the TOMM trial 2 in severe dementia (Teichner & Wagner, 2004), up to 80% in mild to moderate dementia (Walter et al., 2014). Our findings indicate that specificity was 90% in KA patients on the TOMM trial 2, but 75–95% on the indices of the VAT-E. Compared to the earlier reports in dementia, the KA patients performed somewhat better, although compromised. Our findings contrast to the recent observation by Erdodi and Rai (2017) that a single error on the TOMM trial 2 already raises concerns about the credibility of the patient, since 7 out of 20 KA patients (35%) had a single or more errors on the second trial of the TOMM. A single error could therefore also reflect cognitive disorders affecting the sensitivity of PVT test performance. Specifically in

civil litigants cases and forensic settings, false positives are unacceptable (Rogers, 2008), while patients with severe cognitive deficits due to KA or other forms of alcohol-related-brain-damage (ARBD) are relatively overrepresented in such populations (Ekström, Kristiansson, & Björkstén, 2017). Therefore, the findings in our study highlight the importance to interpret PVT performance with caution.

In our study, a logistic regression approach with all the instruments generated a comparable level of prediction as the VAT-E delayed recognition instrument. Importantly, this finding highlights that the KA patients that fail performance on the TOMM due to severe amnesia, also fail performance on the VAT-E. In Table 3 of our manuscript it is confirmed that combining PVTs was not fully successful in eliminating the false positives. The most successful instrument was the profile analysis instrument of the VAT-E that combined scores from multiple resources. In line with Greher and Wodushek (2017), this finding highlights the importance of profile analysis to support the findings of PVTs, instead of relying solely on cutoff data. A different approach that reduced sensitivity to severe amnesia was to combine PVT indices with the SIMS. It could be of practical value in populations with severe cognitive difficulties to investigate both underperformance and over-reporting of symptoms as two important factors of malingering.

Our study included a relatively new instrument to assess both declarative memory functioning and performance validity, called the VAT-E. Earlier evidence indicated that the VAT-E yields very promising results in discriminating patients with Alzheimer's disease, Mild Cognitive Impairment (MCI), and healthy persons that feign memory deficits (Meyer et al., 2017). In our study, not all KA patients were able to pass the VAT-E indices, although only one patient failed the VAT-E Immediate Recognition Trial. All patients were able to pass the profile analysis score. A benefit of applying this novel instrument is that multiple indices to detect underperformance can be combined to more validly assess suboptimal performance. It would be of relevance to study the VAT-E in a larger sample of amnesia patients to investigate whether the sensitivity is comparable in a larger participant group.

In a recent meta-analysis the SIMS was rated as a good test that is well able to distinguish between feigners and nonfeigners of complaints (van Impelen et al., 2014). Earlier results also showed good sensitivity and specificity in a sample of patients with cognitive disorders (Smith & Burger, 1997). In the current study it was found that patients with KA are able to

score sufficiently on the SIMS. None of the KA patients showed exaggeration of symptoms. This implies that the SIMS can be administered safely inpatients that suffer from severe amnesia. In our sample, one healthy comparison reported more symptoms than the cutoff value. Since relatively lower IQ patients were overrepresented in our sample, this could have caused this effect. This notion was earlier made in the SIMS literature, although the relationship in the present study seems to be somewhat stronger, possibly highlighting an overrepresentation of lower IQ's in the present sample. Future research should investigate the SIMS in dementia, since no such studies have been conducted earlier, but are particularly relevant to validate the insensitivity of this test for dementia.

One could argue that the KA patients that underperformed in the present study were not false negatives, but actual underperformers. Neuropsychological tests are only able to measure behavior and not intent, which is essentially unknowable. We are therefore not fully able to rule out the possibility that the KA patients did underperform. However, patients did not have incentives, such as litigations, to underperform in our study, and were actively motivated by the test instructor to engage throughout the experiment. Moreover, the patients in our study were not tested with the PVTs and SVT as part of routine neuropsychological assessment, but were voluntarily participating in this experiment, reducing the motivational effects. Moreover, the testing procedure was restricted to limit the effects of cognitive overload.

Limitations of the present study include the gender composition of the sample, which limits the generalizability of the findings to women. Earlier studies showed that men are overrepresented in KA populations, also reflected in the present study (Kopelman et al., 2009). A second limitation of this study is that both patients and controls were not tested with a larger battery of neuropsychological tests, but rather performed only a few tests. One reason to restrict the number of tests in this study was the relative chance for overloading the patients. Declarative memory tests are particularly stressful for KA patients, because of their generally low performance, and lack of insight into their test performance (Walvoort, van der Heijden, Kessels, & Egger, 2016). Moreover, the samples in the present study were relatively low and, therefore, restricted the statistical power of some of the models. This fact could be explained by the relative scarcity of detoxified KA patients available for neuropsychological testing.

In conclusion, our study shows that patients with severe amnesia are not always able to pass the TOMM, and VAT-E. Since severity of the amnesia was directly related to performance on the PVTs, the performance on those tests should be interpreted with caution in patients with suspected severe amnesia. Combining traditional PVTs was not maximally effective, but including profile analysis resulted in maximum insensitivity to severe amnesia. Also, combining PVTs and the SIMS was effective in eliminating false positives. In clinical use of PVTs, it is therefore relevant to combine levels of evidence to maximize the insensitivity to severe amnesia.

Disclosure statement

Funding

ORCID

Erik Oudman http://orcid.org/0000-0002-4441-0365

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Arlington, TX: American Psychiatric Association.

Bigler, E. D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, 18(04), 632–640. doi:10.1017/S1355617712000252

Bortnik, K. E., Horner, M. D., & Bachman, D. L. (2013). Performance on standard indexes of effort among patients with dementia. *Applied Neuropsychology: Adult*, 20(4), 233–242. doi:10.1080/09084282.2012.695757

Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., … Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity: NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology*, 20, 419–426. doi:10.1016/j.acn.2005.02.002

Dandachi-FitzGerald, B., Ponds, R. W. H. M., Peters, M. J. V., & Merckelbach, H. (2011). Cognitive underperformance and symptom over-reporting in a mixed psychiatric sample. *The Clinical Neuropsychologist*, 25(5), 812–828. doi:10.1080/13854046.2011.583280

Dandachi-FitzGerald, B., Ponds, R. W., & Merten, T. (2013). Symptom validity and neuropsychological assessment: A survey of practices and beliefs of neuropsychologists in six European countries. *Archives of Clinical*

Neuropsychology, *28*(8), 771–783. doi:10.1093/arclin/act073

Dorociak, K. E., Schulze, E. T., Piper, L. E., Molokie, R. E., & Janecek, J. K. (2018). Performance validity testing in a clinical sample of adults with sickle cell disease. *The Clinical Neuropsychologist*, *32*(1), 81–97. doi:10.1080/13854046.2017.1339830

Ekström, A., Kristiansson, M., & Björkstén, K. S. (2017). Dementia and cognitive disorder identified at a forensic psychiatric examination - a study from Sweden. *BMC Geriatrics*, *18*, 219.

Erdodi, L. A., & Rai, J. K. (2017). A single error is one too many: Examining alternative cutoffs on Trial 2 of the TOMM. *Brain Injury*, *31*(10), 1362–1368. doi:10.1080/02699052.2017.1332386

Folstein, M., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State" a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. doi:10.1016/0022-3956(75)90026-6

Greher, M. R., & Wodushek, T. R. (2017). Performance validity testing in neuropsychology: Scientific basis and clinical application—A brief review. *Journal of Psychiatric Practice*, *23*(2), 134–140. doi:10.1097/PRA.0000000000000218

Hampson, N. E., Kemp, S., Coughlan, A. K., Moulin, C. J. A., & Bhakta, B. B. (2014). Effort test performance in clinical acute brain injury, community brain injury, and epilepsy populations. *Applied Neuropsychology: Adult*, *21*(3), 183–194. doi:10.1080/09084282.2013.787425

Iverson, G. L. (2011). *Test of memory malingering. Encyclopedia of Clinical Neuropsychology (2494-2496)*. New York, USA: Springer.

Kopelman, M. D. (2002). Disorders of memory. *Brain : A Journal of Neurology*, *125*(Pt 10), 2152–2190.

Kopelman, M. D., Thomson, A. D., Guerrini, I., & Marshall, E. J. (2009). The Korsakoff syndrome: Clinical aspects, psychology and treatment. *Alcohol & Alcoholism*, *44*, 148–154. doi:10.1093/alcalc/agn118

Lindeboom, J., & Matto, D. (1994). Digit Span and Knoxx Blocks as concentration tests for elderly [in Dutch]. *Tijdschrift Voor Gerontologie en Geriatrie*, *25*, 63–68.

Lindeboom, J., Schmand, B., & Christensen, N. (2003). *Visual Association Test: VAT; Handleiding*. Leiden, Netherlands: PITS.

Liu, Z., Dong, J., Zhao, X., Chen, X., Lippa, S. M., Caroselli, J. S., & Fang, X. (2016). Assessment of feigned cognitive impairment in severe traumatic brain injury patients with the Forced-choice Graphics Memory Test. *Brain and Behavior*, *6*(12), e00593. doi:10.1002/brb3.593

Loring, D. W., Larrabee, G. J., Lee, G. P., & Meador, K. J. (2007). Victoria Symptom Validity Test performance in a heterogenous clinical sample. *Clinical Neuropsychologist*, *21*(3), 522–531. doi:10.1080/13854040600611384

Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist*, *29*(6), 741–776. doi:10.1080/13854046.2015.1087597

McGuire, C., Crawford, S., & Evans, J. J. (2019). Effort testing in dementia assessment: A systematic review. *Archives of Clinical Neuropsychology*, *34*(1):114–131.

Meyer, S., & de Jonghe, J. (2017). *Visuele associatietest - extended: Handleiding*. Amsterdam, Netherlands: Hogrefe.

Meyer, S. R., de Jonghe, J. F., Schmand, B., & Ponds, R. W. (2017). The Visual Association Test-Extended: A cross-sectional study of the performance validity measures. *The Clinical Neuropsychologist*, *31*(4), 798–813. doi:10.1080/13854046.2017.1280181

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1094–1102. doi:10.1076/jcen.24.8.1094.8379

Rogers, R. (Ed.). (2008). *Clinical assessment of malingering and deception*. New York, USA: Guilford Press.

Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S., & Goddard, R. (1986). CAMDEX. A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry*, *149*(6), 698–707.

Rudman, N., Oyebode, J. R., Jones, C. A., & Bentham, P. (2011). An investigation into the validity of effort tests in a working age dementia population. *Aging & Mental Health*, *15*(1), 47–57. doi:10.1080/13607863.2010.508770

Schmand, B. A., Bakker, D., Saan, R. J., & Louman, J. (1991). De Nederlandse Leestest voor Volwassenen: een maat voor het premorbide intelligentieniveau [The Dutch Adult Reading Test: A measure of premorbid intelligence]. *Tijdschrift Voor Gerontologie en Geriatrie*, *22*, 15–19.

Schmand, B., Walstra, G., Lindeboom, J., Teunisse, S., & Jonker, C. (2000). Early detection of Alzheimer's disease using the Cambridge Cognitive Examination (CAMCOG). *Psychological Medicine*, *30*(3), 619–627. doi:10.1017/S0033291799002068

Slick, D. J., Tan, J. E., Strauss, E., Mateer, C. A., Harnadek, M., & Sherman, E. M. (2003). Victoria Symptom Validity Test scores of patients with profound memory impairment: Nonlitigant case studies. *The Clinical Neuropsychologist*, *17*(3), 390–394. doi:10.1076/clin.17.3.390.18090

Sweet, J. J., Benson, L. M., Nelson, N. W. & Moberg, (2015). The American Academy of Clinical Neuropsychology, National Academy of Neuropsychology, and Society for Clinical Neuropsychology (APA Division 40) 2015 TCN Professional Practice and 'Salary Survey': Professional Practices, Beliefs, and Incomes of U.S. Neuropsychologists. *The Clinical Neuropsychologist*, *29*, 1069–1162. doi:10.1080/13854046.2016.1140228

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the American Academy of Psychiatry and the Law Online*, *25*, 183–189.

Teichner, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): Normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology*, *19*(3), 455–464. doi:10.1016/S0887-6177(03)00078-7

Tombaugh, T. N. (1996). *Test of memory malingering: TOMM*. New York, USA/Toronto, Canada: MHS.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment*, 9(3), 260. doi:10.1037/1040-3590.9.3.260

Van Impelen, A., Merckelbach, H., Jelicic, M., & Merten, T. (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. *The Clinical Neuropsychologist*, 28(8), 1336–1365. doi:10.1080/13854046.2014.984763

Moerman - van den Brink, W. G., van Aken, L., Verschuur, E. M. L., Walvoort, S. J. W., Egger, J. I. M., & Kessels, R. P. C. (2019). Executive dysfunction in patients with Korsakoff's syndrome: A theory-driven approach. *Alcohol and Alcoholism*, 54(1):23–29. doi:10.1093/alcalc/agy078.

Walter, J., Morris, J., Swier-Vosnos, A., & Pliskin, N. (2014). Effects of severity of dementia on a symptom validity measure. *Clinical Neuropsychologist*, 28(7), 1197–1208. doi:10.1080/13854046.2014.960454

Walvoort, S. J. W., van der Heijden, P. T., Kessels, R. P. C., & Egger, J. I. M. (2016). Measuring illness insight in patients with alcohol-related cognitive dysfunction using the Q8 questionnaire: A validation study. *Neuropsychiatric Disease and Treatment*, 12, 1609–1615.

Wilson, B. A., Alderman, N., Burguess, P. W., Emslie, H., & Evans, J. J. (1996). *In: Behavioural Assessment of the Dysexecutive Syndrome (BADS)*. Souza Ricardo, O., Schmidt Sergio L., translators. Bury St Edmunds, U.K.: Thames Valley Test Company; Rio de Janeiro, Brazil: Cognição.

Wodushek, T. R., & Greher, M. R. (2017). Performance validity testing in neuropsychology: Methods for measurement development and maximizing diagnostic accuracy. *Journal of Psychiatric Practice*, 23(3), 214–220. doi:10.1097/PRA.0000000000000233