Journal of Neuropsychology (2016), 10, 186–210
© 2014 The British Psychological Society

The British Psychological Society

www.wileyonlinelibrary.com

Dynamic testing of learning potential in adults with cognitive impairments: A systematic review of methodology and predictive value

Hileen Boosman[1], Thamar J. H. Bovend'Eerdt[2], Johanna M. A. Visser-Meily[1], Tanja C. W. Nijboer[1,3] and Caroline M. van Heugten[4]*

[1]Brain Center Rudolf Magnus and Center of Excellence for Rehabilitation Medicine, University Medical Center Utrecht and De Hoogstraat Rehabilitation, The Netherlands
[2]Department of Human Movement Sciences, Maastricht University, The Netherlands
[3]Experimental Psychology, Helmholtz Institute, Utrecht University, The Netherlands
[4]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience and Maastricht University, Department of Neuropsychology and Psychopharmacology, Maastricht University Medical Centre, The Netherlands

Dynamic testing includes procedures that examine the effects of brief training on test performance where pre- to post-training change reflects patients' learning potential. The objective of this systematic review was to provide clinicians and researchers insight into the concept and methodology of dynamic testing and to explore its predictive validity in adult patients with cognitive impairments. The following electronic databases were searched: PubMed, PsychINFO, and Embase/Medline. Of 1141 potentially relevant articles, 24 studies met the inclusion criteria. The mean methodological quality score was 4.6 of 8. Eleven different dynamic tests were used. The majority of studies used dynamic versions of the Wisconsin Card Sorting Test. The training mostly consisted of a combination of performance feedback, reinforcement, expanded instruction, or strategy training. Learning potential was quantified using numerical (post-test score, difference score, gain score, regression residuals) and categorical (groups) indices. In five of six longitudinal studies, learning potential significantly predicted rehabilitation outcome. Three of four studies supported the added value of dynamic testing over conventional testing in predicting rehabilitation outcome. This review provides preliminary support that dynamic tests can provide a valuable addition to conventional tests to assess patients' abilities. Although promising, there was a large variability in methods used for dynamic testing and, therefore, it remains unclear which dynamic testing methods are most appropriate for patients with cognitive impairments. More research is warranted to further evaluate and refine dynamic testing methodology and to further elucidate its predictive validity concerning rehabilitation outcomes relative to other cognitive and functional status indices.

*Correspondence should be addressed to C.M. van Heugten, School for Mental Health and Neuroscience, Department of Psychiatry and Neuropsychology, Maastricht University, P.O. Box 616, Maastricht 6200 MD, The Netherlands (email c.vanheugten@maastrichtuniversity.nl).

DOI:10.1111/jnp.12063

Cognitive impairments are commonly described in sudden onset conditions such as stroke and traumatic brain injury as well as in evolving conditions such as dementia and schizophrenia (Dickinson, Ramsey, & Gold, 2007; Frencham, Fox, & Maybery, 2005; Makin, Turpin, Dennis, & Wardlaw, 2013; Metzler-Baddeley, 2007). In a rehabilitation setting, these patients are generally referred to a neuropsychologist for comprehensive cognitive assessment. Conventional cognitive tests provide information about a patient's baseline performance from which treatment decisions can be made and progress can be monitored (Lezak, Howieson, & Loring, 2004). Neuropsychologists can obtain additional information about patients' abilities by going beyond the standard administration procedures of a cognitive test (Lezak *et al.*, 2004). This can be done by employing a dynamic testing procedure to assess patients' potential to improve cognitive performance.

Dynamic testing is an umbrella term for procedures that examine the effects of a brief training on a person's test performance (Grigorenko & Sternberg, 1998). A commonly used dynamic testing procedure to target cognitive abilities is a one-session pre-test – train – post-test paradigm where a test is administered before and after a brief training (Grigorenko, 2009). During the brief training, the experimenter can, for instance, provide additional instructions or explain compensational strategies. The degree of change between pre- to post-training performance represents a patient's learning potential.

Generally, it is proposed that dynamic tests provide unique information about a person's abilities in addition to the information that is provided by conventional tests (Grigorenko, 2009). It is important to note that dynamic tests were designed to supplement conventional testing procedures rather than replacing them (Grigorenko, 2009). Conventional tests provide valuable information about cognitive deficits that may hamper or facilitate learning, whereas dynamic tests more specifically evaluate patients' potential to learn and improve cognitive performance. Taken together, they can provide a more comprehensive picture of a patient's abilities.

Besides providing additional information about patients' abilities, dynamic tests may also contribute to accurately predicting rehabilitation outcome. Accurate prediction of future achievement is important, as it could guide treatment programmes and identify patients who are in need of individually tailored treatment. The dynamic testing approach has already shown evidence of predictive validity in a review that focused on the use of dynamic testing in academic settings (Caffrey, Fuchs, & Fuchs, 2008). The authors concluded that dynamic tests provided unique information about students' abilities that was not captured by conventional tests; information that contributed to accurately predicting students' future achievement.

The concept of dynamic testing is relatively new for adults with cognitive impairments. Therefore, the objective of this systematic review was to provide clinicians and researchers insight into the concept and methodology of dynamic testing and to explore its added and predictive value in adult patients with cognitive impairments. The following questions were answered:

(1) Which one-session dynamic tests are currently used in adults with cognitive impairments?
(2) Which brief training methods are incorporated into these dynamic tests?
(3) Which computational methods are applied to quantify learning potential?
(4) What is the predictive validity of learning potential concerning rehabilitation outcome?
(5) What is the added value of dynamic tests over conventional tests in predicting rehabilitation outcome?

Methods

Data sources and study selection

The following three electronic databases were searched between January 1990 and May 2014: PubMed, PsychINFO, and Embase/Medline. Search terms were adapted from a previous review on the predictive value of dynamic testing concerning student achievement (Caffrey *et al.*, 2008), and a study on the main features and history of dynamic testing (Grigorenko, 2009). Search terms included dynamic testing, dynamic assessment, learning potential, testing the limits, cognitive plasticity, cognitive modifiability, interactive assessment, mediated learning, mediated assessment, or learntest. The electronic search strategies are shown in Appendix A. Only English language, peer-reviewed journal articles of one-session dynamic tests that include a training phase and target cognitive abilities in adults with acquired cognitive impairments were included. Reviews, dissertations, books, case studies, columns, qualitative studies, neuroimaging studies, psychometric, and methodological evaluations were excluded. Reference lists of full-text eligible articles were screened for relevant articles.

Data extraction

Two authors (HB and TB) independently assessed all studies for inclusion based on the title and abstract. A third author (CvH) was consulted when agreement between the two reviewers was not reached. For all eligible studies, details about the dynamic test, training methods and learning potential indices were extracted. For longitudinal studies that measured rehabilitation outcome on the level of activities and participation, we extracted information about the predictive value of learning potential and the added value of dynamic tests over conventional tests in outcome prediction.

The methodological quality of the included studies was assessed independently by the same two researchers. An 8-point checklist, the Methodological Quality Assessment List (Van Leeuwen, Kraaijeveld, Lindeman, & Post, 2012), was used that yields a total score between 0 (*low quality*) to 8 (*high quality*). This checklist was originally used in spinal cord research. Therefore, item 3 'type of lesion' (paraplegia/tetraplegia) was changed into 'diagnosis'. Studies with a methodological quality score below 3 were excluded from this review. No review protocol was published.

Results

Study selection

In Figure 1, a flow diagram of the search process is presented. The initial search resulted in 1,141 articles. In total, 411 references were duplicates and 695 were excluded based on subject matter. The reference lists of the remaining 35 articles were screened for relevant articles. This yielded three additional articles. The same authors (HB and TB) reviewed the remaining 38 full-text articles and selected studies that were in agreement with the inclusion criteria. The study population characteristics were identical in two articles of the same research group (Fernández-Ballesteros, Zamarrón, & Tárraga, 2005; Fernández-Ballesteros, Zamarrón, Tárraga, & Moya, 2003). The study with the lower methodological quality were excluded for further review (Fernández-Ballesteros *et al.*, 2005). Two other studies also used the same study population (Watzke, Brieger, Kuss, Schöttke, & Wiedl,

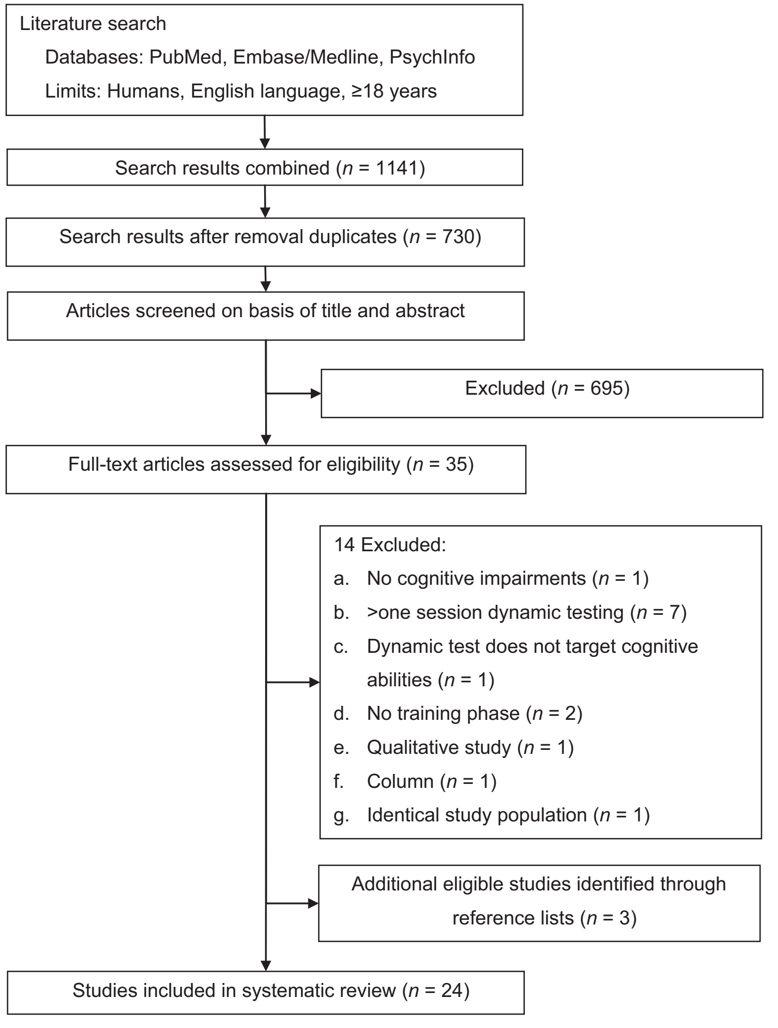

Figure 1. Literature review results flowchart.

2008; Watzke, Brieger, & Wiedl, 2009). Because those studies used different learning potential indices and partially different outcome measures, they were retained for review. The search process resulted in a total of 24 articles, which were included for review.

Study characteristics

Sixteen studies had a cross-sectional design and eight studies had a longitudinal design (Table 1). The methodological quality assessment revealed that on an 8-point scale, the average score was 4.6 (range 3–6). The quality score for 12 of 24 articles was 5 or 6 (see Appendix B). The majority of studies included patients with cognitive impairments due to a psychiatric diagnosis (n = 16). The remaining studies included patients with a neurodegenerative disease (n = 4), acquired brain injury (n = 2) or a combination of patients with a psychiatric or a neurodegenerative disease (n = 2). The included studies determined learning potential for descriptive, diagnostic, or predictive purposes. A

Table 1. Study and participant characteristics

References (country of research)	Test(s)	Population (*n*)*	Mean age (*SD*/range)	% men	Ethnicity, % Caucasian	Treatment programme
Cross-sectional studies						
Fernández-Ballesteros *et al.*, 2003 (Spain)	BEPAD	Healthy controls (100)	73.1 (NR)	49%	NR	NA
		Mild Cognitive Impairment (50)	74.9 (NR)	40%	NR	
		Alzheimer's disease (50)	75.1 (NR)	28%	NR	
Fernández-Ballesteros *et al.*, 2012 (Spain)	BEPAD	Healthy controls 55–89 years (601)	68.8 (6.6)	42%	NR	NA
		Healthy controls >90 years (188)	92.9 (2.5)	36%	NR	
		Mild Cognitive Impairment (57)	76.1 (5.2)	49%	NR	
		Alzheimer's disease (98)	78.2 (5.1)	58%	NR	
Fiszdon *et al.*, 2006 (United States)	CVLT-II	Schizophrenia (40), schizoaffective disorder (10)	42.4 (10.1)	78%	62%	NA
Hake *et al.*, 2007 (United States)	WCST	Schizophrenia (10), schizoaffective disorder (10)	43.5 (NR)	45%	20%	NA
Kolakowsky, 1998 (United States)	BSRT, WMS-R VPA	Experiment 2: Head injury (64)	NR (18–55)	80%	NR	NA
		Experiment 3: Head injury (26)	NR (18–55)	80%	NR	
Kurtz and Wexler, 2006 (United States)	WCST	Schizophrenia, schizoaffective disorder (54)	IEF: 32.7 (10.8) GL: 33.8 (10.1) PL: 36.6 (11.9)	72%	NR	NA
Kurtz *et al.*, 2010 (United States)	WCST	Schizophrenia, schizoaffective disorder (48)	34.7 (12.2)	1%	NR	NA
Rempfer *et al.*, 2006 (United States)	WCST	Bipolar disorder (22), unipolar (major) depression (17), schizophrenia or schizoaffective disorder (21)	43 (10.2)	32%	80%	NA
Rempfer *et al.*, 2012 (United States)	ROCFT	Schizophrenia (45), schizoaffective disorder (36)	41.7 (8.6)	51%	37%	NA
Schreiber and Schneider, 2007 (Germany)	ADAFI	Healthy controls (12)	67.8 (8.3)	33%	NR	NA
		Mild Cognitive Impairment (10)	70.2 (7.1)	30%	NR	

Continued

Table 1. (*Continued*)

References (country of research)	Test(s)	Population (*n*)*	Mean age (*SD*/range)	% men	Ethnicity, % Caucasian	Treatment programme
Uprichard *et al*., 2009 (United Kingdom)	WCST	Closed head trauma (58), stroke (8), anoxic damage (6), cerebral infection (2), missing (3)	30.8 (11.3)	75%	NR	NA
Uttner *et al*., 2010 (Germany)	LMJD	Healthy controls (11) Major depression (11) Alzheimer's disease or at risk (19)	Median 65 (57–85) Median 65 (48–73) Median 72 (49–88)	36% 45% 47%	NR NR NR	NA
Vaskinn *et al*., 2008 (Norway)	WCST	Schizophrenia (26)	32.3 ± 9.3	65%	NR	NA
Vaskinn *et al*., 2009 (Norway)	WCST	Schizophrenia (30)	31.5 (9.6)	67%	100%	NA
Wiedl and Wienöbst, 1999 (Germany)	WCST	Schizophrenia (23)	29.2 (5.2)	52%	NR	NA
Wiedl, Wienöbst *et al*., 2001 (Germany)	WCST	Schizophrenia or schizoaffective disorder (49)	HS: 29.5 (4.3) L: 33.6 (8.1) NL: 34.9 (7.4)	NR	NR	NA
Longitudinal studies						
Calero and Navarro, 2004 (Spain)	AVLT-LP	Healthy controls (101), Mild Cognitive Impairment (102)	74.7 (8.6)	46%	NR	NA
Rempfer *et al*., 2011 (United States)	WCST	Schizophrenia, schizoaffective disorder (9), unipolar (major) depression (8), bipolar disorder (8)	43.8 (NR)	34%	84%	9-session grocery shopping skills training
Sergi *et al*., 2005 (United States)	WCST	Schizophrenia (40), schizoaffective disorder (17)	CIG: 41.4 (10.6) ELG: 41.5 (10.0)	61% 83%	NR NR	1-hr work skills training
Tenhula *et al*., 2007 (United States)	WCST	Schizophrenia (49), schizoaffective disorder (7)	49.4 (6.2)	93%	38%	8-session social skills training

Continued

Table 1. (*Continued*)

References (country of research)	Test(s)	Population (*n*)*	Mean age (*SD*/range)	% men	Ethnicity, % Caucasian	Treatment programme
Watzke *et al*., 2008, 2009 (Germany)	WCST	Schizophrenia, schizoaffective disorder (41)	27.2 (7.4)	61%	NR	1-year vocational rehabilitation programme
Wiedl, Schöttke *et al*., 2001 (Germany)	AVLT, WCST	Sample I: Schizophrenia (29)	30.4 (5.7)	NR	NR	NA
		Sample II: Schizophrenia (33)	34.9 (16.5)	NR	NR	
		Elderly with dementia (50), elderly without dementia (37)	74.8 (8.0)	48%	NR	
Woonings, 2003 (The Netherlands)	M-WCST	Healthy controls (79)	32.7 (7.8)	58%	NR	NA
		Schizophrenia (44)	30.7 (8.2)	86%	NR	8-month rehabilitation programme

Notes. ADAFI, Adaptive Figure Series Learning Test; AVLT, Auditory Verbal Learning Test; AVLT-LP, Auditory Verbal Learning Test of Learning Potential; BEPAD PLPt, VFt, HTPLt, VLMPt, Battery of Learning Potential for Assessing Dementia subtests Position Learning Potential test, Verbal Fluency test, Hanoi Tower Potential Learning test, Verbal Memory Learning Potential test; BSRT, Buschke Selective Reminding test; CVLT-II, California Verbal Learning Test; LMJD, Little mister Jakob drawings; ROCFT, Rey-Osterrieth Complex Figure Test; M-WCST, Modified Wisconsin Card Sorting Test; WCST, Wisconsin Card Sorting Test; WMR-R VPA, Wechsler Memory Scale-Revised Visual Paired Associates; CIG, conventional instruction group; ELG, errorless learning group; GL, good learners; HS, high scorers; IEF, intact executive-function; L, learners; NL, non-learners; PL, poor learners; NR, Not reported; NA, Not applicable.
*Number of participants who completed the dynamic cognitive test.

descriptive purpose was, for example, to explore the association between learning potential and cognitive functioning (Rempfer, Hamera, Brown, & Bothwell, 2006). A diagnostic objective was used only in studies including patients with a neurodegenerative disease. These studies, for example, evaluated whether learning potential can discriminate healthy persons from patients with mild cognitive impairment and Alzheimer's disease (Fernández-Ballesteros *et al.*, 2003). Other studies used learning potential to predict a patient's long-term functioning (e.g., community integration, Uprichard, Kupshik, Pine, & Fletcher, 2009).

Six of eight longitudinal studies measured patients' rehabilitation outcome in terms of activities or participation after a treatment programme. The outcome measures were classified in three domains: community functioning, vocational functioning, or social functioning. The treatment programmes varied from specific skills training (e.g., grocery shopping skills; Rempfer, Brown, & Hamera, 2011), to a comprehensive 8 or 12 months rehabilitation programme (Watzke *et al.*, 2008; Woonings, Appelo, Kluiter, Slooff, & Van den Bosch, 2003).

Measures

Dynamic versions of 11 cognitive tests were employed in the reviewed articles. In Table 2, an overview of all dynamic cognitive tests with their training method and learning potential index is presented. In two studies, dynamic versions of two different cognitive tests were used to assess patients' learning potential (Kolakowsky, 1998; Wiedl, Schöttke, & Calero Garcia, 2001). In the majority of studies (n = 16) dynamic versions of the conventional or modified Wisconsin Card Sorting Test (WCST, Heaton, 1981; M-WCST, Nelson, 1976) were used. Also, all longitudinal studies that evaluated the predictive validity of learning potential used dynamic versions of the WCST (Rempfer *et al.*, 2011; Sergi, Kern, Mintz, & Green, 2005; Tenhula, Strong Kinnaman, & Bellack, 2007; Watzke *et al.*, 2008, 2009) or M-WCST (Woonings *et al.*, 2003). The results for the WCST and M-WCST were combined as these two measures are similar.

Dynamic versions of memory tests were used in seven studies; Auditory Verbal Learning Test (AVLT, Rey, 1964), Auditory Verbal learning Test of Learning Potential (AVLT-LP, Calero & Navarro, 2004), Buschke Selective Reminding Test (BSRT, Buschke, 1973), California Verbal Learning Test-II (CVLT-II, Delis, Kramer, Kaplan, & Ober, 2000), Little Mister Jakob Drawings (LMJD, Uttner *et al.*, 2010), Wechsler Memory Scale-Revised Visual Paired Associates (WMS-R, Wechsler, 1987), and Rey-Osterrieth Complex Figure Test (ROCFT; Rempfer, McDowd, & Brown, 2012). In one study, a dynamic version of a reasoning task was applied; Adaptive Figure Series Learning Test (ADAFI, Guthke & Beckmann, 1995). Two studies used the Battery of Learning Potential for Assessing Dementia (BEPAD, Fernández-Ballesteros *et al.*, 2003, 2012) which includes multiple cognitive domains, namely perception, memory, executive functioning, and language.

Training methods

Training methods that were used most frequently during the brief training were a combination of performance feedback, reinforcement, expanded instruction, or strategy training.

Table 2. Dynamic cognitive tests with their training method(s), and learning potential index

References	Cognitive test(s)	Training method(s)					Learning potential index
		Performance feedback	Reinforcement	Expanded instruction	Strategy training	Other	
Calero and Navarro, 2004	AVLT-LP	✓	✓			✓*	Groups
Fernández-Ballesteros *et al.*, 2003	BEPAD PLPt	✓	✓		✓		Post-test score, difference score
	BEPAD VFt	✓	✓				
	BEPAD VMLPt	✓					
	BEPAD HTPLt					✓†	
Fernández-Ballesteros *et al.*, 2012	BEPAD VMLPt	✓	✓		✓		Post-test score, difference score
Fiszdon *et al.*, 2006	CVLT-II	✓			✓		Groups
Hake *et al.*, 2007	WCST	✓		✓			Groups
Kolakowsky, 1998	BSRT	✓					Groups
	WMS-R VPA				✓		Groups
Kurtz and Wexler, 2006	WCST	✓		✓			Groups
Kurtz *et al.*, 2010	WCST	✓		✓			Gain score
Rempfer *et al.*, 2006	WCST	✓		✓			Groups
Rempfer *et al.*, 2011	WCST	✓		✓			Difference score
Rempfer *et al.*, 2012	ROCFT	✓			✓		Groups
Schreiber and Schneider, 2007	ADAFI	✓				✓†	Post-test score, difference score
Sergi *et al.*, 2005	WCST	✓		✓			Gain score, groups
Tenhula *et al.*, 2007	WCST	✓			✓	✓‡	Regression residuals
Uprichard *et al.*, 2009	WCST	✓		✓		✓§	Groups
Uttner *et al.*, 2010	LMJD				✓		Post-test score
Vaskinn, Sundet, Friis, Simonson *et al.* 2008	WCST	✓		✓			Gain score
Vaskinn *et al.*, 2009	WCST	✓		✓			Gain score, groups
Watzke *et al.*, 2008	WCST	✓		✓			Groups

Continued

Table 2. (*Continued*)

References	Cognitive test(s)	Training method(s)					Learning potential index
		Performance feedback	Reinforcement	Expanded instruction	Strategy training	Other	
Watzke *et al.* 2009	WCST	✓		✓			Post-test score
Wiedl and Wienöbst, 1999	WCST	✓		✓			Groups
Wiedl, Schöttke *et al.*, 2001	AVLT	✓	✓	✓			Post-test score
	WCST	✓		✓			Post-test score
Wiedl, Wienöbst *et al.*, 2001	WCST	✓		✓			Groups
Woonings *et al.*, 2003	M-WCST	✓	✓	✓			Post-test score, groups difference score

Notes. ADAFI, Adaptive Figure Series Learning Test; AVLT, Auditory Verbal Learning Test; AVLT-LP, Auditory Verbal Learning Test of Learning Potential; BEPAD PLPt, VFt, HTPLt, VMLPt, Battery of Learning Potential for Assessing Dementia subtests Position Learning Potential test, Verbal Fluency test, Hanoi Tower Potential Learning test, Verbal Memory Learning Potential test; BSRT, Buschke Selective Reminding test; CVLT-II, California Verbal Learning Test; LMJD, Little mister Jakob drawings; ROCFT, Rey-Osterrieth Complex Figure Test; M-WCST, Modified Wisconsin Card Sorting Test; WCST, Wisconsin Card Sorting Test; WMR-R VPA, Wechsler Memory Scale-Revised Visual Paired Associates.

*Verbalizations aimed at focusing attention on the task.

†Increasing difficulty.

‡Guidance.

§Errorless learning.

Performance feedback

Eight dynamic tests incorporated performance feedback in the brief training. On the ADAFI, error-specific help was given after an incorrect response (Schreiber & Schneider, 2007). For list-learning tests (AVLT, AVLT-LP, BEPAD Verbal Memory Learning Potential Test, BSRT, CVLT-II), feedback mostly consisted of telling the patient the number of words recalled correctly or reminding the patient of words not recalled (Calero & Navarro, 2004; Fernández-Ballesteros *et al.*, 2012; Kolakowsky, 1998; Wiedl, Schöttke *et al.*, 2001). For the ROCFT, patients were corrected after each mistake before continuing with the test and, when needed, patients were given cues for the design elements and organizational sequence during recall (Rempfer *et al.*, 2012). Feedback on the WCST included telling the patient why their choice was correct or incorrect after each card sort (e.g., 'This was wrong, we don't sort for colour, but for form or number'). In contrast, Uprichard *et al.* (2009) provided feedback on the WCST according to an errorless learning approach. After each card sort, patients were asked to say out loud which rule they were using and their count of correct responses. Before starting a new rule, the examiner reminded the patient of the completed and new card sorting rule.

Reinforcement

Four dynamic tests incorporated reinforcement into the brief training. The BEPAD subtests included verbal reinforcement on the patient's performance (e.g., 'you did very well!'; Fernández-Ballesteros *et al.*, 2012). The M-WCST included a monetary reinforcement. Patients received five cents after each correct card sort (Woonings *et al.*, 2003). The type of reinforcement that was used for the AVLT (Wiedl, Schöttke *et al.*, 2001) and AVLT-LP (Calero & Navarro, 2004) was not described.

Expanded instruction

Expanded instruction was given during the brief training of three dynamic tests. The extended, standardized instruction of the AVLT was approximately four times longer than the standard instruction and focused on motivating the patient and ensuring adequate attention (e.g., 'Don't let yourself be distracted by noise or other things'; Wiedl, Schöttke *et al.*, 2001). During the M-WCST training, the examiner gave card-by-card instructions and told the patient of the category shift after ten correct card sorts (Woonings *et al.*, 2003). During the WCST training the sorting rules were explained (e.g., 'There are 3 possible ways to match the cards: You can match the card by colour, by number of the objects, or by shape') and the patient was informed of the rule change after ten consecutively correctly sorted cards (e.g., 'After you get 10 correct in a row, the rule changes; you are no longer matching to colour, you must be matching to the number of objects or to the shape'; Hake, Hamera, & Rempfer, 2007; Kurtz & Wexler, 2006; Kurtz, Jeffrey, & Rose, 2010; Rempfer *et al.*, 2006, 2011; Sergi *et al.*, 2005; Vaskinn, Sundet, Friis, Simonson *et al.* 2008; Vaskinn *et al.* 2009; Watzke *et al.*, 2008, 2009; Wiedl & Wienöbst, 1999; Wiedl, Schöttke *et al.*, 2001; Wiedl, Wienöbst, Schöttke, Green, & Nuechterlein, 2001).

Strategy training

Six dynamic tests incorporated strategy training in the brief training. The BEPAD Position Learning Potential test incorporated verbal and visual strategies separately in order to detect the effect of these different strategies on performance (Fernández-Ballesteros *et al.*, 2003). The BEPAD Verbal Memory Learning Potential test included a cognitive strategy (e.g., 'Perhaps you can group the words'; Fernández-Ballesteros *et al.*, 2012). For the CVLT-II, a semantic memory strategy training was used (Fiszdon *et al.*, 2006). During that training patients were demonstrated that semantic grouping of words increases recall. Patients were given specific instructions on how to group words semantically and were asked to say aloud the semantic groups after recall. The LMJD training included action verbalization (Uttner *et al.*, 2010). The patient was asked to say out loud what was happening on the target pictures. During administration of the WMS-R Visual Paired Associates subtest the patient was instructed to use verbal labelling (Kolakowsky, 1998). In other words, the patients were asked to attach a verbal label to each line drawing. One study incorporated a problem-solving mnemonic in the WCST (Tenhula *et al.*, 2007). The mnemonic included the following steps: (1) identify the problem, (2) identify and select a potential strategy for solving the problem, (3) assess the success of the chosen strategy, and (4) continue to use a successful strategy or revise if the chosen strategy was unsuccessful. For the ROCFT, an organizational strategy was taught. Patients were directed to construct the complex figure in three sequential steps from large structural elements to filling in smaller details. After completion of the drawing, patients were instructed to observe the components and organizational features of the figure (Rempfer *et al.*, 2012).

Computational methods to quantify learning potential

Learning potential was conveyed as a numerical (post-test score, difference score, gain score, regression residuals) or categorical index (groups, e.g., poor learner, strong learner, high-achiever). The computational methods are described below. Six studies used multiple learning potential indices: post-test score and difference score (Fernández-Ballesteros *et al.*, 2003, 2012; Schreiber & Schneider, 2007); gain score and groups (Sergi *et al.*, 2005); and post-test score, difference score, and groups (Woonings *et al.*, 2003).

Post-test score

The post-test score represents the maximum performance a patient can achieve on a cognitive test. A higher post-test score indicates better learning potential. A post-test score was used in seven studies (Fernández-Ballesteros *et al.*, 2003, 2012; Schreiber & Schneider, 2007; Uttner *et al.*, 2010; Watzke *et al.*, 2009; Wiedl, Schöttke *et al.*, 2001).

Difference score

A difference score can be calculated by subtracting the pre-test score from the post-test score. A higher difference score indicates a greater difference between the pre- and post-test. Five studies used a difference score (Fernández-Ballesteros *et al.*, 2003, 2012; Rempfer *et al.*, 2011; Schreiber & Schneider, 2007; Woonings *et al.*, 2003).

Gain score

Gain scores are ratios calculated by dividing actual performance change (i.e., difference score) by potential performance change (Sergi *et al.*, 2005). The lower the score, the lower the learning potential. For example, on the CVLT-II the maximum score a patient can achieve for each trial is 16. This yields the following gain score formula: (trial 5 score – trial 1 score)/(16 – trial 1 score). Thus, the gain score represents the relative change score. Gain scores were calculated in four studies (Kurtz *et al.*, 2010; Sergi *et al.*, 2005; Vaskinn, Sundet, Friis, Simonson *et al.* 2008; Vaskinn *et al.*, 2009).

Regression residuals

This score was calculated by performing a regression analysis in which the pre-test scores was used as a predictor of the post-test scores. The residual scores were used as a measure of learning potential. The higher the score, the greater the difference between the observed and predicted post-test score. One study used regression residuals for the WCST (Tenhula *et al.*, 2007).

Groups

In total, 14 studies divided patients into groups. The following methods were used to do so:

Raw test scores. Patients with a significant increase in the number of words recalled on the BSRT between trials 1 and 4 were classified in a 'learn only' group. Patients who showed an increase in organizational score from trials 4 to 5 on the BSRT, and demonstrated use of a labelling strategy on WMS Visual Paired Associates were classified in a 'learn and benefit group'. Patients who showed adequate scores for semantic labelling strategy were classified in a 'learn, benefit, and transfer group'. This method was used in one study (Kolakowsky, 1998).

Median split. In one study (Sergi *et al.*, 2005), a median split of the change in pre- to post-test T-scores on the WCST was used to classify patients as 'poor learner' or 'strong learner' (Kurtz & Wexler, 2006). Another study used a median split of the gain score (described above) to classify patients in a 'high learning potential' or 'low learning potential' group.

Rasch maps. The difficulty for the different WCST measures (pre- and post-test; e.g., learning to learn, failure to maintain set) was displayed on a 0–100 scale. Patients who scored below the cut-off value for all measures were considered 'non-learners'. Patients who passed some measures were split into 'spontaneous learners' and 'guided learners' based on a midpoint split of 50 on the scale. This method was used in one study (Uprichard *et al.*, 2009).

Algorithm of Schöttke, Bartram, and Wiedl (1993). Linear regression was used to predict the post-test score and a confidence interval was calculated based on the standard error of prediction. The confidence interval was used to determine whether or not the post-test score could be attributed to chance. A post-test score above the confidence interval reflected true change. Patients scoring below the upper limit of the confidence interval were classified as 'poor learners' (also called 'non-retainers' or 'non-learners'). Patients with a post-test score outside the upper limit were considered 'strong learners' (also called 'learners'). Patients with a high pre- and post-test score were classified as 'high achievers' (also called 'high scorers'). Another study used the algorithm to categorize patients in two groups: persons with plasticity or persons without plasticity. This method was used in 12 studies (Calero & Navarro, 2004; Fiszdon *et al.*, 2006; Hake *et al.*, 2007; Rempfer *et al.*, 2006, 2012; Vaskinn, Sundet, Friis, Simonson *et al.* 2008; Vaskinn *et al.*, 2009; Watzke *et al.*, 2008; Wiedl & Wienöbst, 1999; Wiedl, Schöttke *et al.*, 2001; Wiedl, Wienöbst *et al.*, 2001; Woonings *et al.*, 2003).

The predictive and added value of learning potential on rehabilitation outcome

Six of eight longitudinal studies evaluated whether learning potential can predict rehabilitation outcome after a treatment programme. These studies focused on predicting community functioning, vocational, or social functioning. The results are presented in Table 3.

Two studies measured outcome on the level of community functioning (Rempfer *et al.*, 2011; Watzke *et al.*, 2008). Compared to the conventional WCST, the dynamic WCST explained an additional 32% of the variance in shopping skills after a 9-session grocery shopping skills training (Rempfer *et al.*, 2011). Three months after a 1-year vocational rehabilitation programme, non-learners had significantly poorer community functioning compared to learners and high scorers. There were no significant differences between learners and high scorers (Watzke *et al.*, 2008).

Two studies measured outcome on the level of social functioning (Tenhula *et al.*, 2007; Woonings *et al.*, 2003). In the first study, the dynamic WCST was not associated with the change in social functioning between baseline and after an 8-session social skills training. In the other study, it was reported that the conventional M-WCST, and not the dynamic M-WCST, showed a significant, positive association with the change in social functioning between baseline and after an 8-month rehabilitation programme (Woonings *et al.*, 2003). When patients were divided into learner groups, it showed that non-retainers and learners significantly improved to a similar degree, but non-retainers demonstrated significantly lower social functioning compared to learners after the rehabilitation programme (Woonings *et al.*, 2003).

Three studies measured outcome in terms of vocational functioning (Sergi *et al.*, 2005; Watzke *et al.*, 2008, 2009). In one study (Sergi *et al.*, 2005), the authors evaluated the predictive value of the conventional WCST and the dynamic WCST on work skill accuracy and performance 3 months after a 1-hr work skills training. The conventional WCST was a significant predictor of work skill accuracy, whereas the dynamic WCST additionally predicted work skill performance. The conventional WCST explained 6%, and the dynamic WCST explained an additional 13% of the variance in skill accuracy post-training (Sergi *et al.*, 2005). The remaining two studies used the same patient population to evaluate work capabilities and vocational integration after a 1-year vocational rehabilitation programme (Watzke *et al.*, 2008, 2009). One of these studies focused on the predictive value of the conventional WCST and the dynamic WCST on vocational outcomes (Watzke *et al.*, 2009). The other study focused on between-group (non-learners, learners, high scorers) differences in vocational outcome. The dynamic WCST, and not the conventional WCST, was a significant predictor of work capabilities after 6 months of programme attendance (Watzke *et al.*, 2009). At that time, non-learners and learners demonstrated significantly lower work capabilities compared to high scorers.

Table 3. Longitudinal associations between learning potential and rehabilitation outcome

References	Predictor variable	Outcome variable: Measure	Domain	Assessment time	Main results	Results in words
Rempfer *et al.*, 2011	1. WCST pre-test score 2. WCST difference score	TOGSS	Community functioning	Post-treatment	1. $R^2 = 0.27$ 2. $\Delta R^2 = 0.32$***	The conventional WCST explained 27%, and the dynamic WCST difference score explained an additional 32% of the variance in community functioning after a 9-session grocery shopping skills training.
Sergi *et al.*, 2005	1. WCST pre-test score 2. WCST gain score	Work skill accuracy	Vocational functioning	3 months post-treatment	1. $F(1, 52) = 6.87$* $R^2 = 0.06$ 2. $F(1, 52) = 9.69$***, $\Delta R^2 = 0.13$**	The WCST pre-test and gain score were significant predictors of work skill accuracy 3 months after a 1-hr work skills training, where the gain score explained an additional 13% of variance beyond the pre-test score.
	1. WCST pre-test score 2. WCST gain score	Work skill performance	Vocational functioning	3 months post-treatment	1. $F(1, 52) = 2.85$ 2. $F(1, 52) = 9.96$***	The WCST gain score, and not the pre-test score, predicted work skill performance 3 months after a 1-hr work skills training.
Tenhula *et al.*, 2007	WCST regression residuals	MASC	Social functioning	Pre- to post-treatment change	$r = -0.11$ to 0.07	The WCST regression residuals were not associated with the change in social functioning between baseline and after an 8-session social skills training.
Watzke *et al.*, 2008, 2009	WCST group	O-AFP	Vocational functioning	a. After 26 weeks b. Post-treatment	a. NL<HS**, L<HS* b. NL<L,HS*	Compared to learners and/or high achievers, non-learners had lower vocational functioning after 26 weeks, directly after and 3 months after a 1-year vocational rehabilitation programme.
		5-level ordinal scale	Vocational functioning	3 months post-treatment	NL<HS***	
		LFS	Community functioning	3 months post-treatment	NL<L,HS*	Non-learners had poorer community functioning compared to learners and high scorers 3 months after a 1-year vocational rehabilitation programme.

Continued

Table 3. (*Continued*)

References	Predictor variable	Outcome variable			Main results	Results in words
		Measure	Domain	Assessment time		
Watzke *et al.* 2009	1. WCST pre-test score 2. WCST post-test score	O-AFP	Vocational functioning	a. After 6 months b. Post-treatment	1a. $\beta = .312$ 1b. $\beta = .096$ 2a. $\beta = .310$* 2b. $\beta = .148$	The WCST post-test score, and not the pre-test score, was a significant predictor of vocational functioning 6 months after a 1-year vocational rehabilitation programme and not directly after the treatment programme.
	1. WCST pre-test score 2. WCST post-test score	5-level ordinal scale	Vocational functioning	3 months post-treatment	$B = 0.050$ $B = 0.100$**	The post-test score, and not the pre-test score, was a significant predictor of vocational functioning 3 months after a 1-year vocational rehabilitation programme.
Woonings, 2003	1. M-WCST pre-test score 2. M-WCST post-test score 3. M-WCST difference score	REHAB-GB[a]	Social functioning	Pre- to post-treatment change	1. $r = 0.34$* 2. $r = 0.09$ 3. $r = 0.26$	The pre-test score, and not the post-test and difference score, was significantly associated with the change in community functioning between baseline and after an 8-month rehabilitation programme. Compared to learners, non-retainers showed lower community functioning after the programme.
	4. M-WCST group			Post-treatment	4. NR<L*	

Notes. LFS, Level of Functioning Scale; MASC, Maryland Assessment of Social Competence; M-WCST, Modified Wisconsin Card Sorting Test; O-AFP, Osnabrueck Ability to Work Profile; REHAB-GB, Rehabilitation Evaluation Hall and Baker subscale General Behaviour; TOGSS, Test Of Grocery Shopping Skills; WCST, Wisconsin Card Sorting Test; HS, high scorers; L, learners; NL, non-learners; NR, non-retainers.

[a]Higher scores reflects poorer functioning.

*$P \leq .05$, **$P \leq .01$, ***$P \leq .001$.

There were no significant differences between non-learners and learners (Watzke *et al.*, 2008). More important, the dynamic WCST, and not the conventional WCST, significantly predicted patients' vocational integration (Watzke *et al.*, 2009). Regarding learner groups, non-learners demonstrated significantly lower vocational integration compared to high scorers. No significant differences were reported between non-learners and learners, and between learners and high scorers (Watzke *et al.*, 2008).

Discussion

We systematically collected the literature on one-session dynamic testing methods in adults with cognitive impairments and examined the relation between learning potential and rehabilitation outcome. In total, 24 studies were identified describing 11 different dynamic tests that were used to assess learning potential in patients with cognitive impairments. This review provides preliminary support that dynamic tests can provide a valuable addition to conventional tests to predicting rehabilitation outcome. There was, however, a large variability in the methods used for dynamic testing.

Measures

All tests in this review were adaptations of renowned conventional cognitive tests such as the WCST and CVLT-II. The dynamic WCST was used in the majority of studies and was the only test that was administered to patients in all three major diagnostic groups (i.e., acquired brain injury, psychiatric, or neurodegenerative disorders). Of the remaining ten tests, seven were memory tests which reinforces the view that memory is associated with learning (Boosman, Visser-Meily, Winkens, & van Heugten, 2013; Lezak *et al.*, 2004). Memory tests that include repeated administration of a word-list (e.g., CVLT-II) have previously been described as dynamic in nature (Kurtz *et al.*, 2010; Vaskinn, Sundet, Friis, Ueland *et al.*, 2008). However, solely using repetition is in fact unassisted assessment and therefore does not entirely comply with the assumptions of dynamic testing. Repetition of a cognitive test without adding a training phase seems to measure learning effects instead of learning potential.

The BEPAD was unique in the sense that it was the only multi-domain instrument. The BEPAD was developed to test whether learning potential can discriminate healthy persons from persons with Mild Cognitive Impairment and Alzheimer's Disease (Fernández-Ballesteros *et al.*, 2003). The BEPAD subtests were adaptations of existing cognitive tests which had the greatest discriminative power according to experts.

It is important to note that not all cognitive tests can be used for dynamic testing as tests differ in their sensitivity to repeated exposure. Repeated exposure to the same cognitive test may result in practice effects. In particular tests with a single solution are prone to practice effects such as the WCST (Lezak *et al.*, 2004). WCST performance basically depends on discovery of the sort and shift principle. As soon as the sorting principle is discovered or explained, patients are likely to significantly improve their test performance during a second administration of the test (Lezak *et al.*, 2004). The issue of practice effects may be less pronounced for list learning tests as alternative versions can be used (Lezak *et al.*, 2004). Furthermore, some conventional cognitive test may no longer be administered reliably once the patient has performed an adapted dynamic version, especially when explicit instructions are given.

Training methods

Most dynamic tests used a combination of performance feedback, reinforcement, expanded instruction or strategy training. These training methods are commonly used as therapeutic interventions (Hart *et al.*, 2014). In contrast to clinical therapeutic interventions, the brief training during dynamic testing is not intended to provide the patient with information or strategies for use in subsequent rehabilitation or daily life. Dynamic tests merely assess patients' learning potential. The different training methods attempt to induce learning through distinct mechanisms. Providing constructive feedback on performance and giving positive reinforcement and extra instructions may promote learning by enhancing patient's motivation and attention during the task. One of the reviewed studies used a monetary incentive. Although this type of reinforcement may show beneficial effects, its clinical value is questionable in terms of feasibility and ethics. A limitation of providing extra instructions during the training is that the test in some cases becomes a 'one-shot' test depending on the type of instructions given. Strategy training aims at teaching ways to compensate for impairment. Pre- to post-training improvement reflects patients' ability to learn and apply strategies. Strategy training during dynamic testing may not be feasible for all patients. For instance, memory strategy training was dubbed effective in improving recall only in patients with mild memory impairments and not in patients with severe memory impairments (Rees, Marshall, Hartridge, Mackie, & Weiser, 2007). Also, poor performance after strategy training may be the result of the type of strategy training used. There was only one study (Fernández-Ballesteros *et al.*, 2003), in which patients were taught two different strategies separately to detect the effect of these different strategies on test performance.

Learning potential indices

There was a large variability in the computational methods used to quantify learning potential. This variability reflects the discussions in the literature regarding the 'best' or 'preferred' method to quantify learning potential. Several studies have discussed the strengths and limitations of different learning potential indices (Fiszdon & Johannesen, 2010; Waldorf, Wiedl, & Schöttke, 2009; Weingartz, Wiedl, & Watzke, 2008). The post-test score, regression residuals, and learner groups have been favored because of their good stability and validity (Fiszdon & Johannesen, 2010; Weingartz *et al.*, 2008). These learning potential indices were used in the majority of studies in this review. The difference and gain score were also applied in a number of studies, but mostly in conjunction with the post-test score or the group classification. Several drawbacks have been pointed out regarding the use of the difference and gain score to index learning potential. For example, a difference score of zero does not distinguish between a ceiling effect and non-responsiveness (Waldorf *et al.*, 2009), and gain scores can produce disproportionately high or low scores (Weingartz *et al.*, 2008). There are also some issues regarding the use of the post-test score, regression residuals and learner groups to index learning potential. Strictly speaking, the post-test score only measures a patient's maximum performance as it does not assess the amount of learning that has occurred (Fiszdon & Johannesen, 2010). Regression residuals do provide information about the magnitude of change, but are difficult to interpret and are therefore less feasible for use in clinical settings. The group classification provides a clear-cut classification of patients often based on reliable change (Schöttke *et al.*, 1993; Uprichard *et al.*, 2009) at the expense of within-group variation. Hence, the learning potential indices need to be interpreted together to provide a clearer picture of a patient's learning profile, that is, a patient's initial performance, magnitude of change from training and their post-training performance. These scores can also be used to examine within-group variability when using the categorical approach. Any of the learning potential indices viewed in isolation could be misleading.

Learning potential and rehabilitation outcome

The predictive and added value of learning potential was only evaluated for dynamic versions of the WCST and M-WCST and mainly in patients with psychiatric diseases. The results suggest that learning potential can significantly predict rehabilitation outcome in terms of community and vocational functioning (Rempfer *et al.*, 2011; Sergi *et al.*, 2005; Watzke *et al.*, 2008, 2009). Also, support was found for the added value of the dynamic WCST compared to the conventional WCST in predicting community and vocational functioning (Rempfer *et al.*, 2011; Sergi *et al.*, 2005; Watzke *et al.*, 2009).

The predictive value of learning potential on social functioning was only partially supported. One study reported that learning potential was not a significant predictor of social functioning (Tenhula *et al.*, 2007), whereas another study found significant differences in social functioning between learner groups (Woonings *et al.*, 2003). A possible explanation is that social functioning is less reliant on cognitive functioning and thus cognitive learning. Another possible explanation is the use of different learning potential indices. The first study was unique in the use of regression residuals to convey learning potential (Tenhula *et al.*, 2007). The latter study used three different learning potential indices: the post-test and difference score and a group classification (i.e., non-retainers, learners; Woonings *et al.*, 2003). Only the group classification showed significant results. These findings are in line with a recent systematic review and meta-analysis of the dynamic WCST learner groups (Bisoglio, Mervis, & Choi, 2014). The authors of that systematic review reported that poor learning potential was highly predictive of poor response to a psychosocial intervention (OR = 26.44).

Strengths and limitations

One of the strengths of this systematic review is that most search terms were selected from two previous studies (Caffrey *et al.*, 2008; Grigorenko, 2009) and included both specific approaches to dynamic testing (e.g., testing the limits) and major concepts (e.g., cognitive modifiability). We did not limit our search to a specific dynamic testing approach. Furthermore, diagnoses included evolving conditions (neurodegenerative, psychiatric) as well as sudden onset conditions (acquired brain injury). These diagnostic groups are all commonly referred to a neuropsychologist for neuropsychological evaluation and, therefore, provide a good representation of current neuropsychological practice.

A limitation of the studies in this review is that none of the studies had high methodological quality. In particular the internal validity and control of patient drop-out was inadequate. None of the studies tested the validity and reliability of all measurements used, or referred to other studies, which determined the validity and reliability; and none of the studies did a non-response analysis to compare participants and non-participants. Second, between 2010 and 2012 only five studies were published which may indicate a waning interest in dynamic testing. Third, we only included studies in which the whole dynamic testing procedure including the training was done in a 1-day session. This criterion was chosen to minimize the possibility that pre- to post-test improvement can be attributed to other factors than the brief training (e.g., recovery). Last, relatively few studies evaluated the predictive and added value of dynamic testing. For these studies, diagnoses mostly included schizophrenia and schizoaffective disorder. Only one study included patients with acquired brain injury. Interpretation of these results was difficult due to the use of different treatment programmes and outcome measures.

Suggestions for potential clinical applications of dynamic testing

In patients with ABI or a neurodegenerative disease, information about learning potential could be valuable in triage into groups for discharge destination such as in- or outpatient facilities and for the intensity of rehabilitation. For example, patients who are classified as poor learners may need a more intensive or context-dependent, inpatient rehabilitation programme than patients who are classified as high achievers. For the latter group, a less-intensive programme or an outpatient facility may suffice. As most research was performed in psychiatric populations, more research is needed to evaluate the feasibility and predictive value of dynamic testing in patients with ABI or a neurodegenerative disease.

Although this review shows promising results regarding the relation between learning potential and rehabilitation outcome, more research is needed to further evaluate the added value of dynamic testing. It is particularly important to demonstrate that dynamic cognitive tests provide unique information that cannot be captured by conventional cognitive tests, information that can be used to predict individual outcome.

Conclusion

This review provides preliminary support that dynamic tests can provide a valuable addition to conventional tests to assess patients' abilities. Although promising, there was a large variability in methods used for dynamic testing and, therefore, it is unclear which dynamic testing methods are most appropriate for patients with cognitive impairments. More research is warranted to further evaluate and improve dynamic testing methodology and to further elucidate the relation between learning potential and rehabilitation outcome.

Acknowledgements

This study is funded by the National Initiative Brain and Cognition (NIBC). This quick-result project is embedded in the pillar "The Healthy Brain, Program Cognitive Rehabilitation" [grant number 056-11-014]. TCWN supported by the NWO (Netherlands organization for Scientific Research; grant 451-10-013). There were no competing interests.

References

Bisoglio, J., Mervis, J., & Choi, J. (2014). Meta-analysis of "learning potential" on the dynamic Wisconsin Card Sorting Test (D-WCST): Distinct cognitive subgroups with divergent functional outcomes. Abstracts of the 4th Biennial Schizophrenia International Research Conference. *Schizophrenia Research*, *153*, S156. doi:10.1016/S0920-9964(14)70463-3

Boosman, H., Visser-Meily, J. M. A., Winkens, I., & van Heugten, C. M. (2013). Clinicians' views on learning in brain injury rehabilitation. *Brain Injury*, *27*, 685–688. doi:10.3109/02699052.2013.775504

Buschke, H. (1973). Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior*, *12*, 543–550. doi:10.1016/S0022-5371(73)80034-9

Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment. A review. *The Journal of Special Education*, *41*, 254–270. doi:10.1177/0022466907310366

Calero, M. D., & Navarro, E. (2004). Relationship between plasticity, mild cognitive impairment and cognitive decline. *Archives of Clinical Neuropsychology*, *19*, 653–660. doi:10.1016/j.acn.2003.08.008

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test-Second Edition (CVLT-II)*. San Antonio, TX: Psychological Corporation.

Dickinson, D., Ramsey, M. E., & Gold, J. M. (2007). Overlooking the obvious. A meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Archives of General Psychiatry*, *64*, 532–542. doi:10.1001/archpsyc.64.5.532

Fernández-Ballesteros, R., Botella, J., Zamarrón, M. D., Molina, M. A., Cabras, E., Schettini, R., & Tárraga, L. (2012). Cognitive plasticity in normal and pathological aging. *Clinical Interventions in Aging*, *7*, 15–25. doi:10.2147/CIA.S27008

Fernández-Ballesteros, R., Zamarrón, M. D., & Tárraga, L. (2005). Learning potential: A new method for assessing cognitive impairment. *International Psychogeriatrics*, *17*, 119–128. doi:10.1017/S1041610205000992

Fernández-Ballesteros, R., Zamarrón, M. D., Tárraga, L., & Moya, R. (2003). Cognitive plasticity in healthy, mild cognitive impairment (MCI) subjects and Alzheimer's Disease patients: A research project in Spain. *European Psychologist*, *8*, 148–159. doi:10.1027/1016-9040.8.3.148

Fiszdon, J. M., & Johannesen, J. K. (2010). Comparison of computational methods for the evaluation of learning potential in schizophrenia. *Journal of the International Neuropsychological Society*, *16*, 613–620. doi:10.1017/S1355617710000317

Fiszdon, J. M., McClough, J. F., Silverstein, S. M., Bell, M. D., Jaramillo, J. R., & Smith, T. E. (2006). Learning potential as a predictor of readiness for psychosocial rehabilitation in schizophrenia. *Psychiatry Research*, *143*, 159–166. doi:10.1016/j.psychres.2005.09.012

Frencham, K. A., Fox, A. M., & Maybery, M. T. (2005). Neuropsychological studies of mild traumatic brain injury: A meta-analytic review of research since 1995. *Journal of Clinical and Experimental Neuropsychology*, *27*, 334–351. doi:10.1080/13803390490520328

Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, *42*, 111–132. doi:10.1177/0022219408326207

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, *124*, 75–111. doi:10.1037/0033-2909.124.1.75

Guthke, J., & Beckmann, J. F. (1995). *Adaptiver Figurenfolgen-Lerntest (ADAFI) – Manual (Adaptive Figure Series Learning Test – Manual)*. Mödling: Dr. Schuhfried.

Hake, D. L., Hamera, E., & Rempfer, M. (2007). Self-report of cognitive functioning and learning in schizophrenia. *Journal of the American Psychiatric Nurses Association*, *13*, 93–100. doi:10.1177/1078390307303111

Hart, T., Tsaousides, T., Zanca, J. M., Whyte, J., Packel, A., Ferraro, M., & Dijkers, M. P. (2014). Toward a theory-driven classification of rehabilitation treatments. *Archives of Physical Medicine and Rehabilitation*, *95*, S33–S44. doi:10.1016/j.apmr.2013.05.032

Heaton, R. K. (1981). *Wisconsin Card Sorting Test* (WCST). Odessa, FL: Psychological Assessment Resources.

Kolakowsky, S. A. (1998). Assessing learning potential in patients with brain injury: Dynamic assessment. *NeuroRehabilitation*, *11*, 227–238. doi:10.1016/S1053-8135(98)00024-9

Kurtz, M. M., Jeffrey, S. B., & Rose, J. (2010). Elementary neurocognitive function, learning potential and everyday life skills in schizophrenia: What is their relationship? *Schizophrenia Research*, *116*, 280–288. doi:10.1016/j.schres.2009.08.011

Kurtz, M. M., & Wexler, B. E. (2006). Differences in performance and learning proficiency on the Wisconsin Card Sorting Test in schizophrenia: Do they reflect distinct neurocognitive subtypes with distinct functional profiles? *Schizophrenia Research*, *81*, 167–171. doi:10.1016/j.schres.2005.09.003

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment*, 4th ed. New York: Oxford University Press.

Makin, S. D. J., Turpin, S., Dennis, M. S., & Wardlaw, J. M. (2013). Cognitive impairment after lacunar stroke: Systematic review and meta-analysis of incidence, prevalence and comparison with other stroke subtypes. *Journal of Neurology Neurosurgery and Psychiatry*, *84*, 893–900. doi:10.1136/jnnp-2012-303645

Metzler-Baddeley, C. A. (2007). Review of cognitive impairments in dementia with lewy bodies relative to Alzheimer's disease and Parkinson's disease with dementia. *Cortex*, *43*, 583–600. doi:10.1016/S0010-9452(08)70489-1

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, *12*, 313–324. doi:10.1016/S0010-9452(76)80035-4

Rees, L., Marshall, S., Hartridge, C., Mackie, D., & Weiser, M. (2007). Cognitive interventions post acquired brain injury. *Brain Injury*, *21*, 161–200. doi:10.1080/02699050701201813

Rempfer, M., Brown, C., & Hamera, E. (2011). Learning potential as a predictor of skill acquisition in people with serious mental illness. *Psychiatry Research*, *185*, 293–295. doi:10.1016/j.psychres.2009.12.009

Rempfer, M., Hamera, E., Brown, C., & Bothwell, R. J. (2006). Learning proficiency on the Wisconsin Card Sorting Test in people with serious mental illness: What are the cognitive characteristics of good learners? *Schizophrenia Research*, *87*, 316–322. doi:10.1016/j.schres.2006.05.012

Rempfer, M. V., McDowd, J. M., & Brown, C. E. (2012). Assessing learning potential in people with schizophrenia using the Rey Osterrieth Complex Figure Test. *Open Journal of Psychiatry*, *2*, 407–413. doi:10.4236/ojpsych.2012.224057

Rey, A. (1964). *L'examen clinique en psychologie [Clinical examination in psychology]*. Paris: Presse Universitaires de France.

Schöttke, H. H., Bartram, M., & Wiedl, K. H. (1993). Psychometric implications of learning potential assessment: A typological approach. In J. Hamers & K. Sijtsma (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 153–173). Swetz and Zeitlinger: Lisse, The Netherlands.

Schreiber, M., & Schneider, R. (2007). Cognitive plasticity in people at risk for dementia: Optimizing the testing-the-limits-approach. *Aging and Mental Health*, *11*, 75–81. doi:10.1080/13607860600735887

Sergi, M. J., Kern, R. S., Mintz, J., & Green, M. F. (2005). Learning potential and the prediction of work skill acquisition in schizophrenia. *Schizophrenia Bulletin*, *31*, 67–72. doi:10.1093/schbul/sbi007

Tenhula, W. N., Strong Kinnaman, J. E., & Bellack, A. S. (2007). Is learning potential associated with social skills in schizophrenia? *Journal of Rehabilitation Research and Development*, *6*, 827–836. doi:10.1682/JRRD.2006.09.0113

Uprichard, S., Kupshik, G., Pine, K., & Fletcher, B. C. (2009). Dynamic assessment of learning ability improves outcome prediction following acquired brain injury. *Brain Injury*, *23*, 278–290. doi:10.1080/02699050902788444

Uttner, I., Schurig, N., von Arnim, C. A. F., Lange-Asschenfeldt, C., Tumani, H., & Riepe, M. W. (2010). Reduced benefit from mnemonic strategies in early-stage Alzheimer's disease: A brief testing-the-limits paradigm for clinical practice. *Journal of Neurology*, *257*, 1718–1726. doi:10.1007/s00415-010-5610-8

Van Leeuwen, C. M. C., Kraaijeveld, S., Lindeman, E., & Post, M. W. M. (2012). Associations between psychological factors and quality of life ratings in persons with spinal cord injury: A systematic review. *Spinal Cord*, *50*, 174–187. doi:10.1038/sc.2011.120

Vaskinn, A., Sundet, K., Friis, S., Simonsen, C., Birkenaes, A. B., Jónsdóttir, H., . . . Andreassen, O. A. (2008). Emotion perception and learning potential: Mediators between neurocognition and social problem-solving in schizophrenia? *Journal of the International Neuropsychological Society*, *14*, 279–288. doi:10.1017/S1355617708080314

Vaskinn, A., Sundet, K., Friis, S., Ueland, R., Simonsen, C., Birkenaes, A. B., . . . Andreassen, O. A. (2009). Using the Wisconsin Card Sorting Test to assess learning potential in normal IQ schizophrenia: Does it have potential? *Nordic Journal of Psychiatry*, *63*, 405–411. doi:10.1080/08039480902879941

Vaskinn, A., Sundet, K., Friis, S., Ueland, T., Simonsen, C., Birkenaes, A. B., . . . Andreassen, O. A. (2008). Can learning potential in schizophrenia be assessed with the standard CVLT-II? An exploratory study. *Scandinavian Journal of Psychology*, *49*, 179–186. doi:10.1111/j.1467-9450.2008.00627.x

Waldorf, M., Wiedl, K. H., & Schöttke, H. (2009). On the concordance of three reliable change indexes: An analysis applying the dynamic Wisconsin Card Sorting Test. *Journal of Cognitive Education and Psychology*, *8*, 63–80. doi:10.1891/1945-8959.8.1.63

Watzke, S., Brieger, P., Kuss, O., Schöttke, H., & Wiedl, K. H. (2008). A longitudinal study of learning potential and rehabilitation outcome in schizophrenia. *Psychiatric Services*, *59*, 248–255. doi:10.1176/appi.ps.59.3.248

Watzke, S., Brieger, P., & Wiedl, K. H. (2009). Prediction of vocational rehabilitation outcome in schizophrenia: Incremental prognostic validity of learning potential beyond basic cognitive performance. *Journal of Cognitive Education and Psychology*, *8*, 52–62. doi:10.1891/1945-8959.8.1.52

Wechsler, D. (1987). *Wechsler Memory Scale-Revised manual*. San Antonio: The Psychological Corporation.

Weingartz, S., Wiedl, K. H., & Watzke, S. (2008). Dynamic assessment of executive functioning: (how) can we measure change? *Journal of Cognitive Education and Psychology*, *7*, 368–387. doi:10.1891/194589508787724088

Wiedl, K. H., Schöttke, H., & Calero Garcia, M. D. (2001). Dynamic assessment of cognitive rehabilitation potential in schizophrenic persons and in elderly persons with and without dementia. *European Journal of Psychological Assessment*, *17*, 112–119. doi:10.1027/1015-5759.17.2.112

Wiedl, K. H., & Wienöbst, J. (1999). Interindividual differences in cognitive remediation research with schizophrenic patients – indicators of rehabilitation potential? *International Journal of Rehabilitation Research*, *22*, 55–59. doi:10.1097/00004356-199903000-00007

Wiedl, K. H., Wienöbst, J., Schöttke, H. H., Green, M. F., & Nuechterlein, K. H. (2001). Attentional characteristics of schizophrenia patients differing in learning proficiency on the Wisconsin Card Sorting Test. *Schizophrenia Bulletin*, *27*, 687–696. doi:10.1093/oxfordjournals.schbul.a006907

Woonings, F. M. J., Appelo, M. T., Kluiter, H., Slooff, C. J., & van den Bosch, R. J. (2003). Learning (potential) and social functioning in schizophrenia. *Schizophrenia Research*, *59*, 287–296. doi:10.1016/S0920-9964(02)00163-9

Received 23 June 2014; revised version received 2 September 2014

Appendix A: Electronic Search Strategies Used

Database	Search
Pubmed	((((((((dynamic testing[Title/Abstract]) OR dynamic assessment[Title/Abstract]) OR learning potential[Title/Abstract]) OR testing the limits[Title/Abstract]) OR cognitive plasticity[Title/Abstract]) OR cognitive modifiability[Title/Abstract]) OR interactive assessment[Title/Abstract]) OR mediated learning[Title/Abstract]) OR mediated assessment[Title/Abstract]) OR learntest[Title/Abstract] Limits: Humans, English, Adults 19+ years
PsychInfo	(dynamic testing or dynamic assessment or learning potential or testing the limits or cognitive plasticity or cognitive modifiability or interactive assessment or mediated learning or mediated assessment or learntest).ab Limits: Humans, English, adulthood, 18+ years
Embase/Medline	'dynamic testing':ab,ti OR 'dynamic assessment':ab,ti OR 'learning potential':ab,ti OR 'testing-the-limits':ab,ti OR 'cognitive plasticity':ab,ti OR 'cognitive modifiability':ab,ti OR 'interactive assessment':ab,ti OR 'mediated learning':ab,ti OR 'mediated assessment':ab,ti OR 'learntest':ab,ti Limit: Humans, English, Adult, Aged

Appendix B: Methodological Quality Assessment List Scores

References	Item[a,b]								
	1 Internal validity	2 Control of patient drop-out	3 External validity	4 Statistical validity	5 Sample size	6 Control for multicollinearity	7 Control for bias	8 Reporting	1–8 Total
Calero and Navarro, 2004	0	0	0	1	1	NA	0	1	4
Fernández-Ballesteros *et al.*, 2003, 2012	0	0	0	1	1	NA	0	1	4
Fernández-Ballesteros *et al.*, 2012	0	0	0	1	1	0	1	1	4
Fiszdon *et al.*, 2006	0	0	0	1	0	0	1	1	3
Hake *et al.*, 2007	0	0	0	1	1	NA	0	1	4
Kolakowsky, 1998	0	0	0	1	1	0	0	1	3
Kurtz and Wexler, 2006	0	0	1	1	1	NA	1	1	6
Kurtz *et al.*, 2010	0	0	1	1	1	1	1	1	6
Rempfer *et al.*, 2006	0	0	0	1	1	NA	0	1	4
Rempfer *et al.*, 2011	0	0	0	1	1	0	0	1	3
Rempfer *et al.*, 2012	1	0	0	1	1	NA	1	1	6
Schreiber and Schneider, 2007	0	0	0	1	1	NA	1	1	5
Sergi *et al.*, 2005	0	0	1	1	1	1	1	1	6
Tenhula *et al.*, 2007	0	0	1	1	1	NA	0	1	5
Uprichard *et al.*, 2009	0	0	0	1	1	0	1	1	4
Uttner *et al.*, 2010	0	0	1	1	1	NA	0	1	5
Vaskinn, Sundet, Friis, Simonson *et al.* 2008	0	0	1	1	1	1	1	1	6
Vaskinn *et al.*, 2009	0	0	1	1	1	0	1	1	5
Watzke *et al.*, 2008, 2009	0	0	1	1	0	0	1	1	4
Watzke *et al.*, 2009	0	0	1	1	1	0	1	1	5
Wiedl and Wienöbst, 1999	0	0	1	1	1	NA	0	1	5

Continued

Appendix B: (*Continued*)

References	Item[a,b]								
	1 Internal validity	2 Control of patient drop-out	3 External validity	4 Statistical validity	5 Sample size	6 Control for multicollinearity	7 Control for bias	8 Reporting	1–8 Total
Wiedl, Schöttke *et al.*, 2001	0	0	0	1	1	0	1	1	4
Wiedl, Wienöbst *et al.*, 2001	0	0	0	1	1	NA	0	1	4
Woonings *et al.*, 2003	0	0	1	1	1	NA	0	1	5

Notes. [a]0 = negative; 1 = positive; NA = not applicable.
[b]1. Positive if the study tests the validity and reliability of the measures used, or refers to other studies which have established the validity and reliability; 2. Positive if specified how many persons were approached, how many persons participated, and a non-response analysis is done to compare participants and non-participants; 3. Positive if age, gender, diagnosis, and time since diagnosis are specified; 4. Positive if the relationship between a dependent and independent variable is tested for statistical significance; 5. Positive if univariate ratio [n:K] exceeds [20:1] and if multivariate ratio exceeds [n:K] exceeds [10:1]; 6. Positive if specified that multicollinearity between variables has been tested, or if not applicable; 7. Positive if specified that the design accounts for and analyses are corrected for confounders; 8. Positive if purpose is described, results are related to the purpose, and data tables are explained in the results.