Journal of Neuropsychology (2015), 9, 271–298

The British Psychological Society

www.wileyonlinelibrary.com

Reliability and validity of the Leuven Perceptual Organization Screening Test (L-POST)

Kathleen Vancleef[1], Elia Acke[1], Katrien Torfs[2], Nele Demeyere[3], Christophe Lafosse[4], Glyn Humphreys[3], Johan Wagemans[1] and Lee de-Wit[1]*

[1]Department of Experimental Psychology, University of Leuven, Belgium
[2]Institute of Neuroscience, University of Louvain, Louvain-la-neuve, Belgium
[3]Department of Experimental Psychology, University of Oxford, UK
[4]Scientific Unit Rehabilitation Hospital RevArte, Antwerp, Belgium

Neuropsychological tests of visual perception mostly assess high-level processes like object recognition. Object recognition, however, relies on distinct mid-level processes of perceptual organization that are only implicitly tested in classical tests. Furthermore, the psychometric properties of the existing instruments are limited. To fill this gap, the Leuven perceptual organization screening test (L-POST) was developed, in which a wide range of mid-level phenomena are measured in 15 subtests. In this study, we evaluated reliability and validity of the L-POST. Performance on the test is evaluated relative to a norm sample of more than 1,500 healthy control participants. Cronbach's alpha of the norm sample and test–retest correlations for 20 patients provide evidence for adequate reliability of L-POST performance. The convergent and discriminant validity of the test was assessed in 40 brain-damaged patients, whose performance on the L-POST was compared with standard clinical tests of visual perception and other measures of cognitive function. The L-POST showed high sensitivity to visual dysfunction and decreased performance was specific to visual problems. In conclusion, the L-POST is a reliable and valid screening test for perceptual organization. It offers a useful online tool for researchers and clinicians to get a broader overview of the mid-level processes that are preserved or disrupted in a given patient.

Our perception and interaction with the world are primarily visual. In our daily life, visual information is critical at almost every step: Vision guides our movements, enables us to recognize people, places, and objects, and we seek out entertaining visual stimulation. Although vision feels effortless, many complicated processing steps are involved after light is converted into electrical signals on the retina. These processing steps are situated within an interactive modular hierarchy of visual areas (Biederman, 1987; Wagemans, Wichmann, & Op de Beeck, 2005). Within this hierarchy, it is useful to distinguish among three functional stages of processing. First, low-level vision refers to the processing of basic features like orientation, colour, size, length, and location of stimuli in small (local) regions of the visual field. Initially, these attributes are encoded by separate neurons with receptive fields that are only sensitive to very small parts of the incoming stimulation.

*Correspondence should be addressed to Lee de-Wit, Department of Experimental Psychology, University of Leuven, 102 Tiensestraat (box 3711), Leuven 3000, Belgium (email: lee.dewit@ppw.kuleuven.be).

DOI:10.1111/jnp.12050

Second, mid-level processes organize the basic features into more structured units and representations that can then feed into higher level processes. Examples of mid-level processes are perceptual grouping by similarity or proximity, the segregation of figures from backgrounds, the coding of local parts relative to more global wholes, and the construction of shape representations. Third, high-level processes extract meaning from the perceptual structures and make object and face recognition possible by interpreting visual input with cognitive processes involved in the representation of memory and semantic information. It should be noted that although this rough sketch provides a useful framework for understanding vision, a more refined division of labour is undoubtedly required, with many interactions among these three stages (Bar, 2003; Wagemans *et al.*, 2005). Indeed, the final interpretation of features first encoded at low levels is most certainly influenced by computations at the higher stages of processing (Murray, Boyaci, & Kersten, 2006).

In this paper, we focus on the second stage of perceptual organization, or mid-level vision. Perceptual organization is rooted in a rich research tradition starting with the Gestalt psychologists (Wertheimer, 1938; for recent review see Wagemans, Elder, *et al.*, 2012; Wagemans, Feldman, *et al.*, 2012), who were the first to describe basic grouping cues that organize the contents of visual perception. Perceptual grouping refers to the integration of parts into perceptual wholes. One of the most basic of these grouping principles is proximity, which means that neighbouring features are interpreted as more likely to belong to the same object than elements that are further apart. Another grouping principle is common fate, which refers to the grouping of elements that move with the same speed and direction. Grouping is also influenced by collinearity, which means that elements are grouped when they are aligned. Mid-level vision further entails the organization of different groupings into distinct figure and background regions, a process referred to as figure-ground segmentation.

In the last 20 years, developments in cognitive and computational neuroscience have opened up the possibility to study the neural bases of these phenomena and model the mechanisms behind perceptual organization (e.g., Grossberg, Mingolla, & Ross, 1997; Kovács, 1996; Roelfsema, 2006; Spillmann, 1999). Complementary to this, the neuropsychological study of patients with brain damage has also advanced our understanding of perceptual organization. Two famous and important case studies of patients with mid-level vision problems provide a useful illustration. First, after a bilateral lesion to the ventral visual stream, patient DF was unable to discriminate simple geometric shapes and indicate size or orientation (Bridge *et al.*, 2013; Goodale *et al.*, 1994; James, Culham, Humphrey, Milner, & Goodale, 2003; Milner *et al.*, 1991). In addition, she had difficulties in naming objects, in reading, and in copying line drawings. However, her ability to reach for objects was preserved, and when grasping she could correctly adjust her finger grip to the size of the object and rotate her hand in the correct orientation. Nevertheless, when using manual or verbal responses, DF seemed to be insensitive to a range of basic Gestalt grouping cues (Goodale *et al.*, 1994; de-Wit, Kentridge, & Milner, 2009; de-Wit, Kubilius, Op de Beeck, & Wagemans, 2013). DF's case highlights the critical role of the visual ventral stream in many fundamental stages of mid-level visual perception. Another example of the importance of perceptual grouping processes has been documented in patient GK, who perceived his environment in a piecemeal fashion. Unlike DF, he could recognize objects, but could only recognize one object at a time, and in fact could only see multiple stimuli simultaneously if they were connected by some kind of grouping cue (Gilchrist, Humphreys, & Riddoch, 1996).

These case studies highlight the distinct disruptions to mid-level visual processes that can arise following brain damage.

As visual perception is central to human activities, many brain areas contribute to processing visual information and the occipital, parietal, and temporal lobes are all involved in visual perception. It is therefore not surprising that brain damage often affects visual functions. Indeed, while deficits to visual perception might not always be the most obvious among a patient's symptoms, they often provide one of the best predictions of long-term functional impairment (Nys *et al.*, 2005). Deficits can be situated at different levels of visual processing. First, with respect to low-level visual processing, blindsight can occur after damage to the early visual areas in the occipital lobe, in which there is no conscious recognition of stimuli presented in the affected visual field, but the location and movement of the stimuli can correctly be guessed (Weiskrantz, 1990). Cerebral achromatopsia following bilateral damage to ventral occipital areas refers to impaired colour perception (Bouvier & Engel, 2006). Akinetopsia, the loss of movement vision has been reported after bilateral posterior damage (Zihl, von Cramon, & Mai, 1983). Second, deficits in certain mid-level functions that are mediated by lateral areas of the occipital lobe can give rise to apperceptive agnosia, the inability to develop the percept of the structure of an object or a scene which in turn impede object recognition (Milner *et al.*, 1991; Riddoch & Humphreys, 1987). Problems in perceiving more than one object at the time, simultanagnosia, can be present after bilateral injury to parieto-occipital structures (Coslett & Saffran, 1991). Third, more high-level visual deficits following damage to the occipitotemporal border are associative agnosia (impaired access to stored knowledge about objects for vision), prosopagnosia (difficulty in identifying faces), and alexia (inability to read). Besides brain-damaged patients, other patient groups can show altered mid-level visual processing. In autism, for example, changes in perceptual processing have been observed (Dakin & Frith, 2005). Patients with disorganized symptoms of schizophrenia also have deficits in certain aspects of perceptual organization, such as the integration of collinear line segments into a contour (Silverstein & Keane, 2011). Neurodegenerative disorders like Huntington's disease (Lawrence, Watkins, Sahakian, Hodges, & Robbins, 2000) and Alzheimer's disease (Binetti *et al.*, 1998) have been shown to cause problems with visual perception. In sum, deficits in mid-level visual processing frequently occur in a wide range of neurological, psychiatric, and developmental disorders.

Given the contributions of neuropsychological research to our understanding of perceptual organization and the frequency of mid-level deficits in neurological and psychiatric disorders, reliable and valid instruments to measure perceptual organization would be of great value. Unfortunately, there is a lack of well-validated diagnostic tests for adults that independently measure a wide range of mid-level processes. Within the existing instruments, a distinction can be made between tests designed to measure one specific aspect of perceptual organization and more extensive batteries.

With respect to the instruments testing more specific abilities, tests focusing on perceptual grouping are, for instance, the Bender Gestalt Test that makes use of the original stimuli of Wertheimer (Bender, 1938), and the Street Completion Test in which black blobs can be grouped in a meaningful object (e.g., in Eliot & Czarnolewski, 1999). An example of a test measuring figure-ground segmentation is the Poppelreuter–Ghent Test (e.g., in Della Salla, Laiacona, Trivelli, & Spinnler, 1995). Here, participants need to segment overlapping line drawings to recognize the objects. Global–local processing is assessed in the Rey Complex Figure Test (RCFT) (Meyers & Meyers, 1996) and the Embedded Figures Test (e.g., in Barrett, Cabe, & Thornton, 1968). Shape perception can be measured with the Benton Visual Form Discrimination Test (e.g., in Lopez, Charter, Oh, Lazar, & Imperio, 2005) and with the Efron ratio discrimination test (Efron, 1969). The major limitation of these tests is of course that they assess only one aspect of perceptual organization. In addition, many of the tests do not include a clear description of the underlying construct they measure, norms are lacking and validation is not established empirically. Also, the test results are often contaminated with other perceptual and cognitive abilities by including stimuli that are meaningful objects (Street Completion Test, Poppelreuter–Ghent Test) or by relying on motor coordination or memory (RCFT).

Two important batteries of visual functions in adults that measure a wider range of visual processes are the Birmingham Object Recognition Battery (BORB, Riddoch & Humphreys, 1993) and the Visual Object and Space Perception battery (VOSP, Warrington & James, 1991). The BORB is constructed from the hierarchical framework on visual processing that is introduced before. BORB aims to detect impairments of basic processes in visual perception (e.g., apperceptive agnosia) and impairments to vision that are based on access to stored knowledge of objects from vision (e.g., associative agnosia). This twofold goal is reflected in the 14 subtests, of which the first eight evaluate low-level (size, orientation, location and length discrimination) and mid-level visual processes (figure-ground segmentation, viewpoint invariance), while the remaining six tasks assess knowledge of shapes of objects, associative and functional relations of objects and knowledge of object names. The BORB is strongly embedded in the scientific literature on visual perception and provides an indication of deficits at all stages of the visual system. Also, the test is designed to be neglect- and aphasia-friendly. Despite these theoretical qualities and practical advantages, the psychometric properties of the test are somewhat limited. The reports on validity of the BORB have been restricted to reporting descriptive statistics of performance in left- and right-hemisphere patients and reliability indices are unknown. In addition, rather small norm samples are provided, ranging from eight to 39 healthy control subjects per subtest.

The VOSP is designed to detect impaired visual perception that can influence performance on intelligence tests. From this perspective, evaluation of visual perception of objects and space is an essential component of neuropsychological assessment. Preceded by a screening test in which participants have to detect an 'X' in noise, four subtests focus on object perception ('Incomplete letters', 'Silhouettes', 'Shape decision', and 'Progressive silhouettes'), and four other subtests focus on space perception ('Dot counting', 'Position discrimination', 'Number location', and 'Cube analysis'). In the VOSP, no differentiation between levels of visual processing is made. Some subtests explicitly measure object recognition, but implicitly rely on intact low-level vision and mid-level processes like grouping and figure-ground segmentation to structure the image (e.g., 'Silhouettes', 'Progressive silhouettes', 'Shape decision'), while others tap more directly (but not uniquely) into mid-level processes of grouping ('Incomplete letters') and figure-ground segmentation ('Screening test'). Two large and similar norm samples of 200 and 150 patients with extra-cerebral abnormalities are provided with information on the distribution of intelligence and demographic variables. For each subtest, cut-off scores were calculated and average performance of the one of the norm groups is reported (the sample of 200 patients for five subtests; the second sample of 150 patients for the three remaining subtests). Reliability was not evaluated, but validity was investigated by comparing performance of brain-damaged patients with unilateral damage to performance of the healthy control group with a Mann–Whitney *U*-test on each subtest. This comparison indicated significantly worse performance in the right-hemisphere patients than in the left-hemisphere patients and the control group, consistent with the view that visual perception is mainly computed by the right hemisphere (Warrington & James, 1986; Warrington & Taylor, 1973).

In summary, although visual problems and in particular deficits in perceptual organization are important consequences of brain damage and can have a large impact on daily life, neuropsychological measures with adequate psychometric properties are limited. The tests currently available have several limitations with regard to measuring deficits in perceptual organization. The ability to access mid-level processes *per se* is obscured in some tests that rely on high-level processes like object recognition and semantic memory (VOSP, Street Completion Test, Poppelreuter–Ghent Test), spatial attention (VOSP), or motor coordination (RCFT). Other tests measure no more than one or two processes of perceptual organization (Bender Gestalt Test, Street Completion Test, Poppelreuter–Ghent Test, RCFT, Embedded Figure Tests, and Benton Visual Form Discrimination Test). The tests that include a broader range of perceptual processes have some limitations in terms of their psychometric properties, although they are embedded in a strong theoretical framework of visual perception: For example, they only include a small norm sample (BORB), do no report measures of reliability (BORB, VOSP), and few validity studies have been performed (BORB, VOSP). This indicates the need for a neuropsychological instrument that measures a wide range of perceptual organization processes and has adequate psychometric properties.

To meet this need, the Leuven Perceptual Organization Screening Test or L-POST was developed. In 15 subtests, each with five items, several processes of perceptual organization are measured, without relying on long-term memory, language or motor control, which can all be affected with impairments such as amnesia, aphasia, or apraxia. The L-POST is rooted in the rich research tradition of perceptual organization, and stimuli are carefully designed based on recent theoretical and empirical work in vision science and cognitive neuroscience of visual perception. To make the L-POST widely applicable and efficient in use (administering takes only 20–45 min), the test is designed as a screening test of perceptual organization deficits that can be followed up by more in-depth testing of specific visual functions. To reduce cognitive load, the same matching-to-sample task is used in every subtest: Participants have to decide which of the three alternative stimuli at the bottom of the screen is perceived as most similar to the target stimulus at the top of the screen. The test is freely available at www.gestaltrevision.be/tests and is described in detail in Torfs, Vancleef, Lafosse, Wagemans, and de-Wit (2014).

Here, we only give a short description of each subtest (see Torfs *et al.*, 2014 for details). In 'Fine shape discrimination' (subtest 1), participants have to discriminate novel shapes that differ in fine local aspects of the shape, but with similar global properties (Op de Beeck, Baker, DiCarlo, & Kanwisher, 2006; Op de Beeck, Torfs, & Wagemans, 2008). In 'Shape ratio discrimination (Efron)' (subtest 2), the ratio of the width and length of a rectangle has to be judged and the alternative with the same ratio has to be selected. This subtest is an adaptation of the classical assessment of visual form agnosia by Efron (1969). By grouping dots based on proximity in 'Dot lattices' (subtest 3), a general orientation of the pattern can be perceived (Kubovy, Holcombe, & Wagemans, 1998; Kubovy & Wagemans, 1995). Participants have to select the alternative with the same dominant orientation. In 'RFP fragmented outlines' (subtest 4), grouping of line elements to perceive a shape is necessary. In this task, participants have to use the principle of good continuation (Koffka, 1922) and no distractor background elements are present. In 'RFP contour integration' (subtest 5), participants have to group collinear Gabor elements to perceive a non-familiar shape and segment it from a background of random oriented Gabor elements (similar to Machilsen & Wagemans, 2011). In 'RFP texture surfaces' (subtest 6), texture patterns of similarly oriented lines make segregation of an irregular shape from the background possible (e.g., Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999). To solve 'Global motion detection' (subtest 7), detection of the movement direction of coherently moving dots (with a common fate) in random-dot kinematograms is necessary (Williams & Sekuler, 1984). In 'Kinetic object segmentation' (subtest 8), participants have to group coherently moving Gabor elements inside and outside of a simple geometric shape. The distinct motion patterns create a kinetic boundary and give rise to the percept of a shape surrounded by a background (e.g., Segaert, Nygård, & Wagemans, 2009). In 'Biological motion' (subtest 9), participants have to use coherent motion signals and structure-from-motion to perceive the representation of a walking human in a pattern of moving dots and to discriminate this from spatially scrambled coherently moving dots (Johansson, 1973; Troje, 2002). In 'Dot counting' (subtest 10), participants have to recognize the number of dots presented during short flashes of 200 ms. The short presentation time requires grouping (and probably subitizing; Kaufman, Lord, Reese, & Volkmann, 1949; Trick, 2008) of the dots and renders serial counting impossible. In 'Figure-ground segmentation' (subtest 11), a correct interpretation of the figure-ground relations is necessary to amodally complete the shape that is occluded by circular discs (Nakayama, Shimojo, & Silverman, 1989). In 'Embedded figure detection' (subtest 12), participants have to select the complex geometric line pattern in which a simple target pattern is embedded. This subtest taps into part–whole processing for which patients will have to focus on local elements to solve the task (Witkin, 1962). In 'Recognition of missing part' (subtest 13), participants have to select the alternative in which the same detail is omitted as in the meaningful target object. This tasks taps into part–whole encoding and a more local focus is necessary to solve the task. In 'Recognition of objects in isolation' (subtest 14), recognition of everyday objects on a white background is assessed. This task serves as a control task for 'Recognition of objects in a scene' (subtest 15) in which the same objects are presented in a natural scene. By comparing performance on both tasks, we assess scene segmentation in the context of object recognition. See Figure 1 for an example of the items for each subtest.

In the current study, we investigated the psychometric properties of the L-POST. First, we evaluated reliability in a healthy norm sample and in a sample of brain-damaged patients. Second, we investigated the construct validity of the L-POST by three complementary procedures: (1) We investigated the association between performance on the L-POST and relevant (e.g., brain damage) and irrelevant (e.g., sex) biographical variables. (2) Next, convergent validity was assessed by comparing L-POST performance with performance on similar neuropsychological measures, while (3) the correlations with performance on test of other cognitive domains like memory, attention or executive functions gave insights into the discriminant validity. Third, we studied the internal structure of the L-POST by means of confirmatory factor analyses. In these analyses, we assessed if the intended structure of sets of subtests tapping into a specific process of perceptual organization was represented in the data.

Method

Participants

To create our norm sample, we selected a subsample out of 2,678 participants without a history of brain lesion who completed the L-POST; this selection was based on a number of criteria. First, the data from 77 participants who took the test a second time were

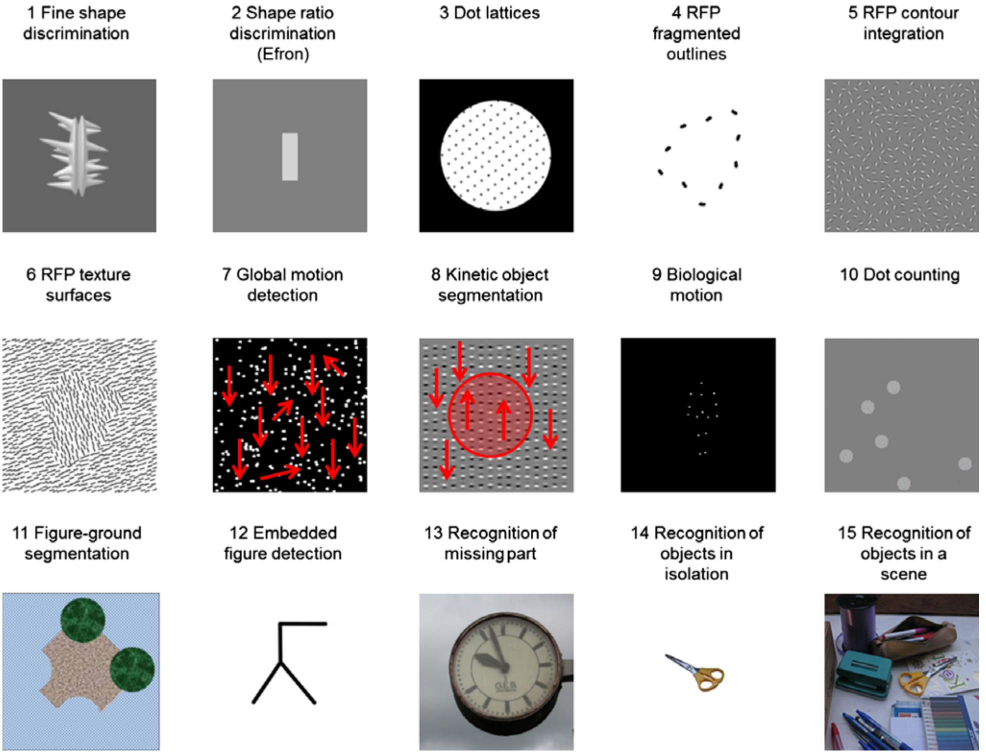

Figure 1. Stimulus examples of each of the 15 subtests of the L-POST. Red arrows indicate the motion directions of the elements and are only added for illustrative purpose.

excluded. We also excluded data of participants with a visual disorder that could not be corrected by glasses or contact lenses (85 participants) and problems in motivation, communication, general intellectual impairment, general depression, or other related problems (278 participants). Also, data from suboptimal test conditions were not included in the norm sample, for example, in case of technical problems (slow internet connection, problems in loading the images, 121 participants) or a small screen size that could not fit all stimuli at once (27 participants). At the end of the test, participants were asked if they filled-in the test seriously and could indicate if they were interrupted during the test on a 7-point scale ranging from ‘not at all’ to ‘continuously’. Data from participants who reported an interruption level of 3 or higher (162 participants) or did not take the test seriously (42 participants) were also excluded from the norm sample. Last, only data of adults were included in the norm sample (278 children were excluded). These inclusion and exclusion criteria resulted in a norm sample of 1,567 participants. Participants’ ages ranged from 18 until 88 with more younger than older participants (Figure 2); 41% of our norm sample finished high school, 22% finished higher vocational education, 27% obtained a Bachelor or Master degree, 3% had a PhD, and 7% reported ‘Other’ as highest education level. In our sample, 37% of the participants were male and 63% were female; 87% of the participants were right-handed, 10% left-handed and 3% ambidextrous. Testing conditions varied over subjects: 26% of the participants completed the L-POST at home without supervision, 59% at home under supervision of an undergraduate student, and 15% in a collective testing session in a

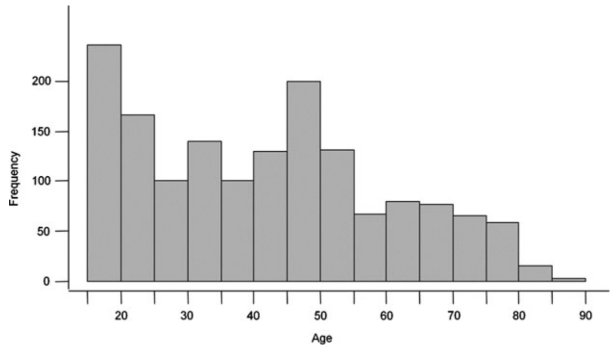

Figure 2. Distribution of ages of the healthy norm sample.

computer lab in the presence of a researcher. These collective sessions included 183 psychology students who received course credit for taking part in the study. This student sample causes the overrepresentation of females between 18 and 25. All participants who completed the L-POST received no further instructions than those provided on the website.

In addition to the healthy control norm sample, data were collected from 58 brain-damaged patients to evaluate construct validity of the L-POST. The same inclusion and exclusion criteria were used as for the healthy control sample, which resulted in a patient sample of 52 participants. Demographic details of the patients and lesion description are provided in Appendix S1. Twenty patients were recruited at the University of Oxford (United Kingdom). They completed an English version of the L-POST and other cognitive functions were assessed with Birmingham Cognitive Screen (BCoS) to evaluate convergent validity and discriminant validity (see Table 1). All tests were administered during home visits by a researcher. The other 32 patients were tested at RevArte Rehabilitation Hospital (Belgium) with a Dutch version of the L-POST; 12 of them only participated in the test–retest reliability study, the other 20 patients participated in the reliability and validity studies and therefore completed additional tests. For the patients from RevArte, we also measured visual acuity with the Freiburg Visual Acuity Test (Bach, 1996). All 32 patients showed 'near normal' vision following the criteria of the World Health Organization (International Classification of Diseases 10): 24 patients showed normal vision (0.8–1.6 decimal Visual Acuity [dec VA]), in seven patients 'mild vision loss' (0.32–0.63 dec VA) was indicated, and one patient had above normal vision (>1.6 dec VA). To evaluate convergent and discriminant validity, comparable Dutch and English neuropsychological measures were selected. However, some subtests of BCoS are language-dependent and are currently not available in Dutch. These subtests were replaced by equivalent neuropsychological tests that are translated in Dutch (Table 1). To evaluate test–retest reliability, 20 RevArte patients completed the L-POST a second time (eight of them also participated in the validity study). To assess the reliability, a 2-week time interval was adopted to minimize both the effect of spontaneous recovery in brain-damaged patients (that increases over time) and the effect of memory on the test scores (that decreases over time). In one patient, the time interval was 3 weeks. Note that not all patients completed all neuropsychological tests because of fatigue, motor problems, language problems, etc. (see Tables 3 and 4).

Table 1. Overview of neuropsychological measures for convergent and discriminant validity

Type of validity	Oxford validation patient sample	RevArte validation patient sample
Convergent validity	BORB size matching	BORB size matching
	BORB overlapping figures	BORB overlapping figures
	VOSP screening test	VOSP screening test
	VOSP incomplete letters	VOSP incomplete letters
	VOSP dot counting	VOSP dot counting
	BCoS complex figure copy	Rey Complex Figure Test: copy trial
Discriminant validity		
Attention	BCoS apple cancellation	BCoS apple cancellation
	BCoS visual extinction	BCoS visual extinction
	BCoS tactile extinction	BCoS tactile extinction
Executive functions	BCoS rule finding and concept switching	Behavioural Assessment of the Dysexecutive Syndrome
	BCoS auditory attention task	
Language	BCoS picture naming	AAT confrontation naming
	BCoS sentence construction	AAT token test
	BCoS sentence reading	AAT repetition
	BCoS reading non-words	AAT written language
	BCoS writing words and non-words	AAT comprehension
	BCoS instruction comprehension	
Memory	BCoS orientation	Rivermead Behavioural Memory Test
	BCoS story recall and recognition	
	BCoS task recognition	
Number skills	BCoS number/price/time reading	
	BCoS number/price writing	
	BCoS calculation	
Praxis	BCoS multi-step object use	BCoS gesture imitation
	BCoS gesture production	
	BCoS gesture recognition	
	BCoS gesture imitation	

Note. AAT, Aachen Aphasia test; BCoS, Birmingham Cognitive Screen; BORB, Birmingham Object Recognition Battery; VOSP, Visual Object and Space Perception battery.

Instruments

Table 1 shows an overview of the tests administered for convergent and discriminant validity in each patient sample.

Leuven Perceptual Organization Screening Test

In the L-POST, different processes in perceptual organization were measured in 15 subtests (Figure 1). For each subtest, five items were presented. To reduce cognitive load and increase the consistency across the subtests, a matching-to-sample task was used for all test items: Participants had to indicate the alternative that is most similar to the target stimulus. Subtests were administered in a random order, with the restriction that the 'Recognition of objects in a scene' subtest always preceded the 'Recognition of objects in isolation' subtest. Two general scoring systems were used. First, we calculated an absolute total score based on the number of correct items (maximum is 75). Second, for the patients, a relative total score was obtained by comparing performance on each subtest with the norm sample. We counted the number of subtests in which a participant scored below the 10th percentile of the norm sample. By calculating both an absolute and a relative score, we could detect patients who have a very selective deficit that is only reflected in two or three subtests. These patients would obtain a relatively high absoulte total score, while their relative deficit is indicated by the number of failed subtests. In addition to the general scoring, for each subtest, the number of correct items was taken as a measure (maximum is 5). As 'Objects in isolation' is a control condition for 'Object in a scene', we calculated the difference score between both subtests to reflect the additional difficulty of recognition when objects are presented in a scene.

Birmingham Object Recognition Battery

The BORB measures aspects of visual perception in apperceptive tests and associative tests (Riddoch & Humphreys, 1993). Only the apperceptive subtests 'Size matching' and 'Overlapping figures' were administered in our patient samples. In 'Size matching', patients had to judge if the size of two circular discs was the same or different in 30 items with an equal number of same and different items. The number of correct answers was taken as an accuracy measure. In the 'Overlapping figures' subtest, figure-ground segmentation was assessed: Patients had to name figures that were presented in isolation, as non-overlapping pairs, as overlapping pairs, as non-overlapping triplets, or as overlapping triplets. The figures could be letters or line drawings of recognizable objects. The line drawings were only presented in isolation or in pairs, not in triplets. In this subtest, the performance measure was the ratio between the reaction time for naming the non-overlapping figures and the reaction time for naming the overlapping figures. No reliability estimates or empirical assessments of validity are available for the BORB.

Visual Object and Space Perception battery

Tasks of the VOSP can be divided into two categories: object recognition and spatial relations (Warrington & James, 1991). Before taking these tasks, participants are screened with a shape detection task in which they have to detect an 'X' in noise. This screening task was administered in the patients, together with 'Incomplete letters' and 'Dot counting'. In 'Incomplete letters', patients had to recognize letters that were degraded by removing square-shaped parts of the letters. As is evident from the name, in the 'Dot counting' task, patients had to count the number of dots on the page ranging from five to nine dots. We scored the number of correct responses in each subtest. Measures of reliability are not provided in the manual. Validity is assessed by comparing brain-damaged patients with a healthy norm sample. No difference between both groups was observed in the screening test, but a substantial number of patients scored below the 5th percentile cut-off in the 'Incomplete Letters' and 'Dot counting'.

Birmingham Cognitive Screen

The BCoS was developed for the comprehensive and efficient screening of a range of cognitive abilities (Humphreys, Bickerton, Samson, & Riddoch, 2012). The assessment provides a profile of the cognitive abilities and challenges of the patient.

Attention and executive function were measured in ‘Auditory attention task’, ‘Rule finding and concept switching’, ‘Apple cancellation’, ‘Visual extinction’, and ‘Tactile extinction’. The ‘Auditory attention task’ provided measures of selective attention, sustained attention, and working memory. In the task, participants had to respond to three auditory presented target words and suppress responses to three closely related distractors. In ‘Rule finding and concept switching’, participants had to predict the movement of a black marker in a grid that moved in a lawful way. Along the task, the abstract movement rule was changed. Accuracy was taken as a measure of this task. Egocentric and allocentric neglect was assessed in ‘Apple cancellation’ in which participants had to cross the full apples on an A4 sheet of paper along distractors of apples with a gap at either the left or right side. Extinction in the visual and tactile domain was measured in ‘Visual extinction’ and ‘Tactile extinction’, respectively. In these tasks, the left and right visual field or left and right hand were stimulated in unilateral and bilateral trials. Extinction was indicated by an inability to detect stimuli in bilateral trials, while detection in unilateral trials was normal. We compared the largest performance difference between both types of trials (left or right) with performance on the L-POST.

Language was measured with ‘Picture naming’, ‘Sentence construction’, ‘Sentence reading’, ‘Reading nonwords’, ‘Writing words and nonwords’, and ‘Instruction comprehension’. In ‘Picture naming’, participants had to name grey-shaded hand drawings of living and non-living objects to asses visual object recognition and access to semantic/conceptual knowledge. ‘Sentence construction’ measured semantic and syntactic processes by asking participants to describe what the person on a photograph was doing using two provided words. In ‘Sentence reading’, participants had to read two sentences to test for the ability to read different word classes. Phonological procedures of reading were measured in ‘Reading nonwords’. Phonological and lexical spelling was assessed in ‘Writing words and nonwords’. ‘Instruction comprehension’ on four target tasks was assessed by a clinical judgement of the examiner.

The tasks that assessed memory are ‘Orientation’, ‘Story recall and recognition’, and ‘Task recognition’. In ‘Orientation’, access to personal information and orientation in time and space is evaluated with open questions. Long- and short-term memory is assessed in ‘Story recall and recognition’. This task provides different scores for memory encoding, retrieval, and consolidation. Because this differentiation is not relevant for our purpose, these scores were summed to obtain one overall score on this task. Unintentional memory was assessed in ‘Task recognition’ in which the participant was presented with four objects of which one had been used in a previous task.

Number skills were measured in ‘Number/price/time reading’, ‘Number/price writing’ and ‘Calculation’. Functional measures of the processing of numbers in everyday situations were measured in ‘Number/price/time reading’ and ‘Number/price writing’. Coding of numbers and basic number-processing operations were measured in ‘Calculation’.

The five tasks measuring praxis were ‘Complex figure copy’, ‘Multi-step object use’, ‘Gesture production’, ‘Gesture recognition’, and ‘Imitation’. In ‘Complex figure copy’ constructional apraxia, organization of the figure, and visual neglect were assessed. In ‘Multi-step object use’, participants had to perform a sequence of actions with multiple objects. Meaningful and meaningless gestures were included in ‘Gesture production’, ‘Gesture recognition’, and ‘Imitation’.

Reliability was investigated in test–retest study with 20 control participants and 17 patients with chronic neurological damage. Correlations, Wilcoxon signed rank tests, *t*-tests, and percentages of agreement between the scores of both sessions were obtained for each task and indicated overall high test–retest reliability. Inter-rater reliability on the tasks that require qualitative judgements of performance ranged from .94 to .99. A factor analysis on the task scores and correlations between performance on BCoS tasks and on established external measures of specific cognitive functions indicated high construct validity.

Rivermead Behavioural Memory Test

The Rivermead Behavioural Memory Test (RBMT) assesses memory capacities in brain-damaged patients (Wilson, Cockburn, & Baddeley, 1989). Eleven tasks measured short-term memory in immediate recall tasks (remembering a story, a short route, delivering a message), long-term memory in delayed recall tasks (recalling the same story, the short route, a name, line drawings, faces), orientation in time and space, prospective memory skills (delivering a message, asking for a hidden belonging, remembering an appointment), and the ability to learn new information (learning new faces, a story, a route, a name). All tasks were related to everyday life. Screening scores were calculated for each patient by taking the total score over all tasks. This score reflected general memory problems. Parallel-form reliability of the RBMT was on average .77 for the four forms, test–retest reliability of the screening score was .78. Uniform positive correlations with other memory tests, observations of memory lapses and a questionnaire of everyday life memory problems indicated good convergent validity of the battery (Wilson, Cockburn, Baddeley, & Hiorns, 1991).

Rey Complex Figure Test

In the RCFT, patients are asked to reproduce a complex line drawing, first by copying, later from memory immediately after the copy trial and after a 30-min delay (Meyers & Meyers, 1996). A recognition trial with parts of the complex figure could also be administered. The test assesses visuoconstructional abilities and visuospatial memory. As we were more interested in visuospatial abilities, patients only completed the copy trial and not the memory tasks. We scored the accuracy and placement of the individual elements of the drawing. In addition, an Organizational Strategy Score was given (Anderson, Anderson, & Garth, 2001). Inter-rater reliability of the accuracy and placement scores ranged from .93 to .99, and 100% agreement of the clinical interpretation was found on two occasions of testing. The test had adequate convergent, discriminant, and factorial validity.

Aachen Aphasia Test

Language disorders were assessed with a Dutch version of the Aachen Aphasia Test (AAT; 'Akense Afasie Test'; Graetz, De Bleser, & Willmes, 1992). This test measures aphasia symptoms in speaking, listening, reading, and writing in six subtests. It allows for differentiation of types of aphasia (e.g., global aphasia, Broca's aphasia, or Wernicke's aphasia). The different subtests are 'Spontaneous speech', 'Token test', 'Repetition', 'Written language', 'Confrontation naming', and 'Comprehension'. With the exception of the first subtest, all subtests were administered in the RevArte patient sample. In the Dutch version, reliability is expressed by Cronbach's alpha between .88 (for 'Comprehension') and .99 (for 'Repetition'). The internal structure of the test was evaluated by a cluster analysis and showed evidence for the differentiation between the subtests and increasing item difficulty within each subtest. In addition, successful discrimination between aphasia and non-aphasia patients (in 93% of the cases) and between types of aphasia (in 88% of the cases) indicates good discriminant validity.

Behavioural Assessment of the Dysexecutive Syndrome

With a battery of seven tests, Behavioural Assessment of the Dysexecutive Syndrome (BADS) measures executive functions generally associated with the frontal lobe (Wilson, Alderman, Burgess, Emslie, & Evans, 1996). The subtests are designed to represent real-life situations to increase the ecological validity of the test. The test requires patients to plan, initiate, monitor, and adjust behaviour. Subtests are 'Rule shift cards', 'Action programme', 'Key search', 'Temporal judgement', 'Zoo map', 'Modified six elements,' and a 'Dysexecutive questionnaire'. The profile score that is the sum of the subtest scores was taken as our measure of executive functioning. Although inter-rater reliability was high (between .88 and 1), test–retest correlation was ranging between −.08 and .71 because of strong learning effects. The BADS showed construct validity in differentiating between brain-damaged patients and the control group and is comparable to other established tests of executive functions (Norris & Tate, 2000).

Data-analyses

Reliability

Reliability was evaluated both in our norm sample and in our patient sample. In the norm sample, we calculated Cronbach's alpha for the total score and at the level of the subtests. Cronbach's alpha indicates a lower bound of reliability at population level and is based on covariances between the subtests, the variance of the overall score, and the number of subtests. In the patient sample, we calculated test–retest correlation for the total score and at the level of subtests. Pearson correlations were calculated for the continuous variables 'total score' and 'number of failed subtests'. At the level of subtests, polychoric correlations were calculated to account for the limited range of possible values (0–5). Given the small sample ($n = 20$), we performed non-parametric permutation tests to determine the p-values.

External structure: Difference between samples

To evaluate construct validity, we calculated performance differences between groups on relevant (brain damage) and irrelevant (age, setting, sex and education level) dimensions. Because of the large number of subjects in the norm sample, effect sizes give a better indication of the effect at population level than mere p-values.

Convergent and discriminant validity

Convergent and discriminant validity were evaluated by comparing the performance on the L-POST with performance on other perceptual neuropsychological measures in the patient sample. We calculated Pearson correlations between performance on the L-POST (total score and number of failed subtests) and performance on convergent and discriminant tests if test scores are continuous and can at least take 15 different values. Polychoric correlations with L-POST performance measures were calculated for categorical variables with an underlying continuous distribution or if test scores have a limited range (<15). If the sample size was smaller than 30, permutation tests with 500 random permutations of the data were applied to calculate the *p*-values.

Internal structure: Factor analyses

To study the internal structure of the L-POST, factor analyses were performed on the data of the healthy norm sample. As the L-POST was developed from a theoretical framework of perceptual organization, two alternative hypotheses about the factor structure could be generated and confirmatory factor analyses were applied. The goodness of fit of our two-factor structure models was explored in a random half of the norm sample ($n = 783$). Subsequently, the preferred model was cross-validated in the independent other half of the norm sample ($n = 784$). Loadings of the subtests on the relevant factor were to be freely estimated, while the loadings on the other factors were fixed to zero. All factors were allowed to correlate. As the subtest scores had a limited range (1–5) and data were highly skewed, the scores were treated as ordinal, and factor analyses were based on the polychoric correlation matrix. For the same reason, diagonally weighted least squares were used to estimate the model parameters. However, to compute robust standard errors, and a mean- and variance-adjusted test statistic, the full weight matrix is used. For least squares estimators, no likelihood is provided, which implies alternative models cannot be compared by likelihood statistics (e.g., AIC or BIC). Therefore, model fit was evaluated and compared on robust estimates of fit indices χ^2, comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA). For cut-off criteria of these fit indices, we followed the guidelines of Hu and Bentler (1999), who suggest .95 for CFI and TLI and a cut-off close to .06 for RMSEA to conclude a relatively good fit between the hypothesized model and the observed data. As a significant χ^2 relative to the degrees of freedom indicates a difference between the observed and the model implied variance–covariance matrices, a good fit is associated with a non-significant value. However, because *p*-values of χ^2 decrease with sample size, careful interpretation is advisable. Analyses were performed with lavaan (Rosseel, 2012), a structural equation modelling package for the statistical software R (R Development Core Team, 2011).

The first model was constructed based on simple stimulus characteristics to provide a benchmark to compare our second model, which attempted to capture the underlying mid-level processes engaged in the various sub-tests. The factors included in the first model were ‘Motion’ and ‘Colour’. Motion-related subtests were ‘Kinetic object segmentation’, ‘Biological motion’, ‘Global motion detection’, and ‘Dot counting’. Subtests with coloured stimuli were ‘Figure-ground segmentation’, ‘Recognition of missing part’, and the difference score between the subtests ‘Recognition of objects in isolation’ and ‘Recognition of objects in a scene’. For the other subtests, the loadings on both factors were fixed at zero. Our second model included factors representing perceptual processes. The first factor *Perceptual grouping* refers to the process of combining elements into a whole. This included the subtest ‘RFP fragmented outline’ and ‘RFP contour integration’ in which elements have to be grouped to perceive the shapes, ‘Dot lattices’, in which dots have to be grouped based on proximity to perceived the direction, ‘Biological motion’ and ‘Global motion detection’, in which grouping coherently moving dots gives rise to a perceived moving direction or the interpretation of a walking figure, and ‘Dot counting’, in which dots have to be rapidly grouped in configurations that are easy recognizable (e.g., the configuration of dots on a dice) to estimate the number of dots. The second factor *Figure-ground segmentation* refers to the process of segmenting two surfaces and dissociating a figure from a background, which was the central process in the subtest 'Figure-ground segmentation', and in 'Recognition of objects in a scene' (difference score). These tests were loading on this factor. Also in 'Kinetic object segmentation' and in 'RFP texture surfaces' segmentation of two surfaces is necessary to perceive the shape. Third, attention to specific parts of the stimulus was relevant in 'Recognition of missing part' and 'Embedded figure detection' which were related to the *Parts in whole* factor. In the last factor, *Shape discrimination* fine judgements about the shape are central to the task. The subtests loading on this factor were 'Fine shape discrimination' and 'Shape ratio discrimination'.

Results

Reliability

We evaluated reliability in both the norm sample and the patient sample. In our healthy norm sample, moderate reliability was suggested by Cronbach's alpha of .76 (based on total score) and a mean intersubtest correlation of .18. Reliability estimates of the subtests are presented in Table 2. Because these estimates are based on only five items per subtest, low Cronbach's alpha could be expected. In the patient sample, test–retest reliability was evaluated. We observed good test–retest reliability on the total score ($r = .77$, $p < .001$) and on the number of failed subtests ($r = .77$, $p < .001$). Polychoric correlations and associated p-values for each subtest are presented in Table 2.

As we observed satisfactory reliability at the general performance measures 'total score' and 'number of failed subtests' (for patients only), but not at the level of individual subtests, only general performance scores were used in further analyses.

Table 2. Cronbach's alpha and test–retest reliability for each subtest

Subtests	Cronbach's alpha in healthy norm group (n = 1,567)	Test–retest reliability in brain-damaged patients (n = 20)
1. Fine shape discrimination	.42	.53 (.07)
2. Dot lattices	.52	**.54 (.04)**
3. RFP contour	.49	.46 (.06)
4. Shape ratio discrimination (Efron)	.25	.56 (.16)
5. Figure-ground segmentation	.42	**.60 (<.01)**
6. Kinetic object segmentation	.54	**.77 (<.01)**
7. RFP texture surface	.39	**.67 (.02)**
8. RFP fragmented outline	.35	−.09 (.82)
9. Biological motion	.64	**.51 (.04)**
10. Recognition of missing part	.50	.50 (.08)
11 Global motion detection	**.80**	**.94 (<.01)**
12. Dot counting	.57	**.56 (.03)**
14. Embedded figure detection	.46	−.03 (.98)
13–15. Object recognition in a scene	.37	**.90 (<.01)**

Note. A bold value indicates a Cronbach's alpha higher than .7 or a test–retest correlation that is significantly different from 0.
The p-values of the polychoric correlations are shown between brackets.

Validity

External structure: Difference between samples

As expected, we observed a large and significant difference between brain-damaged patients (n = 40) and healthy control participants (n = 1,567) in the total score on the L-POST ($M_{\text{patients}} = 59.5$, $M_{\text{controls}} = 70.7$, $t(40) = 39.29$, $p < .001$, $d = 1.26$). In addition, we evaluated if the total score on the L-POST was associated with less relevant demographical variables in our norm sample. We only observed a small effect of age on the L-POST scores, $F(1, 1{,}565) = 85.95$, $p < .001$, adjusted $R^2 = .05$, explaining 5% of the variance in the L-POST scores: L-POST scores decrease a little with age. Also, the effect of testing condition is small, $F(2, 1{,}564) = 16.85$, $p < .001$, $\omega^2 = .02$, explaining only 2% of the variance: participants who took the test at home in absence of a research assistant obtained a somewhat higher score than participants who were supervised by a research assistant either at home or in the lab. In addition, male participants performed slightly better than female participants, $t(1{,}294) = 5.80$, $p < .001$, $d = 0.30$, and we observed a trend for better performance with higher education levels, $F(4, 1565) = 26.01$, $p < .001$, $\omega^2 = .06$. The substantial significance levels of these small effects is a consequence of our large norm sample.

Convergent validity

Moderate but significant correlations were observed between the total score on the L-POST and 'Size matching' in the BORB and two of the three measures of 'Overlapping figures' of the BORB (see Table 3). In the third measure of 'Overlapping figures', the

Table 3. Correlations between performance on the L-POST and other neuropsychological tests of visual perception

			L-POST total score		L-POST number of failed subtests	
Neuropsychological measure	Sample	*n*	*r*	*p*	*r*	*p*
BORB						
Size matching	O+R	40	.54	**<.001**	−.48	**.002**
Overlapping figures: paired letters	O+R	39	.35	.031	−.24	.137
Overlapping figures: letter triplets	O+R	39	.43	**.006**	−.37	.021
Overlapping figures: paired line drawings	O+R	39	.43	**.006**	−.44	**.005**
VOSP						
Shape detection	O+R	39	.03	.840	−.07	.662
Incomplete letters	O+R	39	.62	**<.001**	−.49	**.001**
Dot counting	O+R	39	.36	.025	−.34	.036
Complex figure tests						
Total score RCFT	R	20	.75	**<.001**	−.67	**.001**
Organizational strategy score RCFT	R	20	.40	.108	−.37	.148
BCoS complex figure copy	O	19	.78	**<.001**	−.79	**<.001**

Note. Bold values indicate significant effects after Bonferroni–Holm correction for multiple comparisons at an initial alpha of .05. We applies this correction separately for our two patient samples.
O = Oxford patient sample; R = RevArte patient sample; RCFT = Rey Complex Figure test; BCoS = Birmingham Cognitive Screen; BORB = Birmingham Object Recognition Battery; VOSP = Visual Object and Space Perception battery.

Table 4. Correlations between performance on the L-POST and neuropsychological tests of other cognitive functions

			L-POST total score		L-POST number of failed subtests	
Neuropsychological measure	Sample	*n*	*r*	*p*	*r*	*p*
Spatial attention						
BCoS apple cancellation: egocentric neglect	O+R	39	.02	.880	.03	.810
BCoS apple cancellation: allocentric neglect	O+R	39	.22	.180	−.29	.100
BCoS visual extinction	O+R	40	−.56	.012	.59	**<.001**
BCoS tactile extinction	O+R	39	−.43	.024	.46	.028
Executive functions						
BCoS rule finding and concept switching	O	19	.61	.004	−.57	.024
BCoS auditory attention task: selective attention	O	20	.43	.080	−.38	.120
BCoS auditory attention task: sustained attention	O	20	−.31	.200	.16	.590
BCoS auditory attention task: working memory	O	20	.76	.360	−.74	.020
BADS profile score	R	19	.64	**<.001**	−.67	**<.001**
Language						
BCoS picture naming	O	19	.38	.100	−.28	.328
BCoS sentence construction	O	19	−.30	.408	.35	.276
BCoS sentence reading	O	18	−.03	.900	−.05	.812
BCoS reading non-words	O	18	.08	.792	−.07	.784
BCoS writing words and non-words	O	17	−.25	.420	.22	.492
BCoS instruction comprehension	O	19	.03	.920	.11	.752
AAT confrontation naming	R	17	.63	**<.001**	−.58	.012
AAT token test	R	17	−.59	.020	.56	.030
AAT repetition	R	17	.79	**<.001**	−.73	**<.001**
AAT written language	R	17	.64	**.004**	−.63	.012
AAT comprehension	R	17	.59	.020	−.56	.044
Memory						
BCoS orientation	O	19	.04	.936	.04	.824
BCoS story recall and recognition	O	19	.11	.640	−.09	.740
BCoS task recognition	O	19	.32	.208	−.16	.564
RBMT	R	20	.63	.008	−.55	.016
Number skills						
BCoS number/price/time reading	O	19	.12	.680	−.20	.420
BCoS number/price writing	O	19	.11	.728	−.15	.680
BCoS calculation	O	19	.49	.076	−.42	.100
Praxis						
BCoS multi-step object use	O	19	.00	.920	.12	.744
BCoS gesture production	O	19	.28	.300	−.17	.532
BCoS gesture recognition	O	19	.35	.188	−.20	.468
BCoS gesture imitation	O+R	39	.37	.032	−.35	.060

Note. Bold values indicate significant effects after Bonferroni–Holm correction for multiple comparisons at an initial alpha of .05. We applied this correction separately for our two patient samples.
O = Oxford patient sample; R = RevArte patient sample; RBMT = Rivermead Behavioural Memory Test; BCoS = Birmingham Cognitive Screen; BADS = Behavioural Assessment of the Dysexecutive Syndrome; AAT = Aachen Aphasia Test.

Table 5. Standardized factor loadings for each subtest of the perceptual processes model

Subtest	Perceptual grouping	Figure-ground segmentation	Parts in whole	Shape discrimination
4 RFP fragmented outline	.71			
5 RFP contour integration	.65			
10 Dot counting	.60			
3 Dot lattices	.53			
7 Global motion detection	.52			
9 Biological motion	.48			
5 Kinetic object segmentation		.78		
6 RFP texture surfaces		.74		
11 Figure-ground segmentation		.66		
14–15 Object recognition in a scene		.30		
13 Recognition of missing part			.60	
12 Embedded figure detection			.55	
1 Fine shape discrimination				.76
2 Shape ratio discrimination (Efron)				.53

correlation was marginally significant. The ‘Shape detection’ task of VOSP did not correlate with performance on the L-POST. The subtest ‘Incomplete letters’ and ‘Dot counting’ did correlate with the total score on the L-POST, although the correlation with ‘Dot counting’ was only marginally significant. Regarding copying a complex figure, we observed a strong correlation between L-POST performance and the score on BCoS ‘Complex figure copy’, L-POST performance, and the total score on RCFT, but not between L-POST performance and the Organizational Strategy Score of RCFT. Correlations between the convergent tests and the number of failed subtests in the L-POST followed the same pattern.

Discriminant validity

Discriminant validity was evaluated in six cognitive domains: spatial attention, executive functions, language, memory, number skills, and praxis (see Table 4). In the spatial attention domain, we found no significant correlations between performance on the L-POST and measures of neglect. A significant correlation was observed between visual extinction and the number of failed subtests in the L-POST. Regarding executive functions, we found a significant correlation between L-POST performance and the BADS profile score. For the other executive functions, no significant correlations were observed. In the language domain, we found no evidence for strong correlations with performance on the L-POST and BCoS language tests. However, three subtests of the AAT: ‘Confrontation naming’, ‘Written language’, and ‘Repetition’ correlated strongly with performance on the L-POST. No significant correlations with the other subtests of the AAT were found. Also, no significant correlations were observed with measures of memory or number skills and praxis.

Internal structure

Descriptive statistics on L-POST subtest performance, including means and variance–covariance matrix, are presented in Appendix S2. Our confirmatory factor analyses indicated inadequate fit indices for the stimulus characteristics model, $\chi^2(83) = 2{,}178.73$,

$p < .001$; CFI = .08; TLI = −0.01; RMSEA = .18, but good fit indices for the perceptual processes model, $\chi^2(71) = 86.25$, $p = .11$; CFI = .99; TLI = .99; RMSEA = .02, 95% CI (0, .03). Based on our theoretical considerations and exploration of the correlation structure in the data, we adapted the perceptual processes model as our final model (see Table 5). Subsequently, this model was cross-validated in the second half of the norm sample, which indicated a good fit, $\chi^2(71) = 73.30$, $p = .40$; CFI = 1; TLI = 1; RMSEA = .01, 95% CI (0, .02).

Discussion

Reliability

At the level of general performance on the L-POST, test–retest correlations indicated good reliability in the patient group (.77 on the number of failed subtest and .77 for the total score). In the healthy norm sample, Cronbach's alpha provided a lower bound of reliability of .76 (total score). Given the diverse range of mid-level processes measured in the L-POST, which was confirmed by the four factor structure, we did not expect a high Cronbach's alpha on the total score and the observed value showed acceptable reliability. Also, the calculation of Cronbach's alpha was based on the scores of only 15 subtests and lengthening the test to 20 subtests would already increase Cronbach's alpha to .81 as estimated by the Spearman-Brown formula (and to .85 with 25 subtests). We concluded that at a group level, conclusions about L-POST performance can be derived with confidence. In addition, according to the criteria described in Nunnally and Bernstein (1994), the L-POST has sufficient reliability for the classification of individual patients, although the administration of complementary measures of visual perception would be advisable when important decisions about the patient have to be made (e.g., driving capability).

At the level of the subtests, not enough items (only five per subtest) are provided to result in high values of Cronbach's alpha. However, for eight of the subtests, we found test–retest correlations significantly different from zero. Given these results, hypotheses about the impaired processes of perceptual organization can be generated based on the subtest scores. However, it is advisable not to derive conclusions about deficits in specific perceptual processes on the basis of the subtest scores, at least not at an individual level. In follow-up testing with measures designed for specific aspects of perceptual organization (see overview in the introduction), these hypotheses can be verified. This approach corresponds to the aims of the L-POST, to provide a screening for potential mid-level deficits that will require more detailed follow-up testing to identify the specific problems. For this purpose, the L-POST shows adequate reliability.

Validity

A first indication for construct validity was provided by the evaluation of differences on L-POST performance between samples based on biographical variables. These analyses showed a strong effect of brain damage on L-POST performance, while no major effects of sex, age, education level, and test setting were observed. The extra factors did have a statistically reliable effect (because of the large samples), but the effect sizes were small (explaining between 2% and 6% of the variance). However, previously, effects of ageing on perceptual organization have been observed (e.g., Gilmore, Tobias, & Royer, 1985;

Humphrey & Kramer, 1999) and for the Embedded Figures Test, a positive association with education and males has been reported (Goodenough, 1976).

Second, a comparison of L-POST performance with performance on other measures of visual perception showed good convergent validity. We found substantial correlations between L-POST performance and related tests of visual perception: 'Size matching' and 'Overlapping figures' of the BORB, 'Incomplete letters' of the VOSP and the copy condition of the RCFT and of the complex figure in BCoS, indicating good convergent validity of the L-POST. Moreover, correlations are not high enough to conclude that the L-POST measures identical constructs to these subtests of the BORB and the VOSP, which would make the L-POST redundant. Surprisingly, very small correlations were found with the shape detection screening task of the VOSP, even though this task does not rely heavily on object recognition or reading ability. In addition, a rather moderate correlation with overlapping paired letters was observed. Also, the correlation with the Organizational Strategy Score of RCFT was lower than expected. This finding could be attributed to the limited variance in the Organizational Strategy Scores of the RCFT: 16 of 20 patients obtained a score of four or five and qualitative differences between these two levels are small. Last, 'Dot counting' of the VOSP correlated only moderately with L-POST performance even though the L-POST includes a similar task. However, in the L-POST, the dots are presented in flashes of 200 ms and reappear at random locations, while in the VOSP, static presentations of the dots are used. Therefore, it is possible in the VOSP to serially count the dots on the page, while in the L-POST, rapid grouping of the dots is essential to instantaneously recognize the number of dots presented and maintaining information over flashes is necessary. The distinction between counting and subitizing might be relevant to this difference in performance (Kaufman *et al.*, 1949) and different brain areas are involved in both processes (Demeyere, Lestou, & Humphreys, 2010). As a more general remark on the moderate correlations, we found with the BORB and VOSP, it must be emphasized that these latter tests do not purely measure perceptual organization, and reports on construct validity are limited. Therefore, the moderate correlations with L-POST performance could be attributed to other cognitive processes like spatial perception and object recognition that influence performance in the BORB and VOSP more than in the L-POST.

Third, discriminant validity was evaluated by comparing L-POST performance with neuropsychological measures of spatial attention, executive functions, memory, language, number skills, and praxis. We mostly found small correlations between deficits in these functions and performance on the L-POST, indicating that the L-POST is specific for visual problems and that performance is not highly influenced by other cognitive impairments.

One exception is the relatively high correlation between BCoS visual extinction and the number of failed subtests. This indicates that patients with strong visual extinction perform worse on the L-POST and that their impairments are not restricted to one or two subtests or subprocesses of perceptual organization (as there is only a substantial correlation with the number of failed subtests and not with the total score). However, for consistency, all extinction patients in our sample completed the standard version of the L-POST and not a neglect-compatible version, in which the response stimuli are presented in one vertical column instead of one horizontal row. A vertically aligned (neglect-compatible) version of the L-POST is available online. To decide if this version is required for a given patient, a short neglect test was also developed in which the patient is asked to click on four grey squares. These squares are presented at the same positions as the actual stimuli in the L-POST. In case visual extinction is suspected, it is therefore advisable to perform the custom-made neglect test, before proceeding with the standard or the neglect-friendly version of the L-POST. Note that we observed moderate correlations with BCoS tactile extinction, but this correlation can be attributed to the association between tactile and visual extinction. Indeed, when calculating the partial correlation between L-POST performance and BCoS tactical extinction, while controlling for visual extinction, the correlation decreased to $-.07$.

Concerning executive functions, we found that BADS scores correlated with L-POST performance. However, several subtests of the BADS rely on intact visual perception like the cards that are used in 'Rule shift cards', the 'Zoo map', the 'Action program' or the 'Key search'. Indeed, the BCoS measures of executive functions that do not include visual tasks do not correlate with L-POST performance.

The correlations with languages tests are more diverse: We found no correlations with BCoS measures of language, but we did find significant correlations between L-POST performance and AAT language measures. This striking result might be attributed to the different patient sample. While the Oxford patients who completed BCoS were at a chronic stage (at least 6 months post-stroke), the RevArte patient sample included sub-acute patients. In sub-acute patients, poor performance on language test does not imply specific language problems. Patients with very low working memory capacity will also score low on the language comprehension tasks especially when these sentences are increasing in complexity. *Post-hoc* correlations indeed show an association between executive functions as measured with BADS and two of three AAT language measures (AAT confrontation naming: $r = .68$, $p = .08$; AAT repetition: $r = .45$, $p = .26$; AAT written language: $r = .74$, $p = .04$). With respect to the other cognitive domains, memory, number skills and praxis, no correlations with L-POST performance were found.

Given the correlations between L-POST performance and visual extinction (BCoS), executive functions (BADS), and language (AAT), we also calculated correlations between the existing visual perception tests (BORB and VOSP) and visual extinction (BCoS), executive functions (BADS) and language (AAT) to compare discriminant validity of the L-POST with discriminant validity of BORB and VOSP. These correlations (see Appendix S3) indicate that also the subtests of the BOBR and VOSP show rather unexpected correlations with other cognitive measures. This result highlights that the scores of the L-POST (and BORB and VOSP) need to be interpreted with care and within the broader clinical profile of each patient.

The results on convergent and discriminant validity indicate that the L-POST can confidently identify patients with visual deficits, although performance on the L-POST can also be influenced by problems in other cognitive domains (extinction, executive functions, and language). Therefore, while good performance on the L-POST indicates adequate visual perception, poor performance on the L-POST should always be considered in terms of the broader clinical profile of each patient as revealed by other clinical tests.

A factor analysis confirmed the theoretically implied structure of the L-POST. Four perceptual factors were identified: *Perceptual grouping, Figure-ground segmentation, Parts in wholes,* and *Shape discrimination*. The theoretical foundation of the L-POST with stimuli derived from recent studies in cognitive neuroscience has led to a consistent structure of the L-POST as is indicated by the excellent fit indices. Cross-validation of the model in an independent sample indicated sample invariance of the factor structure.

Characteristics of norm sample

In the context of this study, we collected data from more than 2,000 healthy control participants of which about 1,500 were included in the norm sample after strict selection. Because the L-POST is a free online test, the norm sample is increasing every day with an average of 75 participants every week. This large norm sample is representative for the population as it includes participants of a wide age range and of all educational levels. Women aged between 18 and 25 years are slightly overrepresented because of participant recruitment among psychology students. In the online interface, one could also select specific norm groups to compare to a particular patient for specific test or research purposes.

Applications

In the current study, we have shown that the L-POST is a reliable and valid screening instrument for deficits in mid-level visual perception. It is a useful instrument in both clinical evaluation of an individual patient and in neuropsychological research. In the clinic, the L-POST can complement the BORB and the VOSP in neuropsychological assessment of brain-damaged patients. It is a brief, free and easy-to-use instrument, with practical tools for clinicians like automatic scoring, a one-page printable report of the results, and user-adjustable norm samples. In addition, neuropsychological research can benefit from the L-POST. The L-POST can be used to identify patients with problems in perceptual organization. The specific deficits of these patients can then be studied in depth with patient-tailored experiments and correlated with the lesions to gain more insights into the neurological underpinnings and the mechanisms of visual perception. It can also be used as a control condition to test for other deficits which imply a mid-level or high-level deficit (e.g., to rule out a general problem of configural processing in prosopagnosia patients). For researchers, we offer the possibility to download the raw data of their patients for advanced analyses.

Besides brain-damaged patients, other patient populations might benefit from a screening instrument for perceptual organization. First, impairments of perceptual organization in schizophrenia have been documented as the sixties and have been selected as a consistent biomarker of the disorganized symptoms of the disorder that have been related to poor pre-morbid social functioning, and poor prognosis (Silverstein & Keane, 2011). The L-POST could provide a sensitive measure of the disorganized subtype of schizophrenia (Schenkel, Spaulding, & Silverstein, 2005; Uhlhaas, Phillips, Mitchell, & Silverstein, 2006). Also, because impaired perceptual organization has typically only been observed in chronic patients, the L-POST could give an index of illness progression (Silverstein & Keane, 2011; Uhlhaas *et al.*, 2006). Second, there is already evidence that Alzheimer's disease causes deficits in visual perception, but this has only been established in experimental studies. Whilst only 57% of the patients revealed a clear deficit on object recognition, all patients had a deficit in figure-ground analysis (Mendez, Mendez, Martin, Smyth, & Whitehouse, 1990). Also, some aspects of perception like the ability to judge the ratio between the width and height of a simple shape are intact in these patients, while other aspects of perception seem to be impaired (Tippett, Blackwood, & Farah, 2003). The L-POST could provide a first indication of which processes are spared and which are impaired in an Alzheimer patient. Third, the L-POST can be a valuable instrument in developmental psychology to study normal development of perceptual organization (Pisella *et al.*, 2012) and altered perceptual organization in development disorders like autism spectrum disorder. It has been shown that children with autism spectrum disorders have an advantage over typical developing children in tasks that ask for a local focus, like the Embedded Figures Task (Frith, 1989). To assess this altered balance in local-global processing, the L-POST might be useful. Besides autism spectrum disorders, also patients with eating disorders show weak central coherence (Lopez, Tchanturia, Stahl, & Treasure, 2008a,b; Lopez, Tchanturia, Stahl, Booth, *et al.*, 2008; Southgate, Tchanturia, & Treasure, 2005). Potentially, this is a cognitive style that is not limited to visual processing, but is also manifest in other behaviour, for instance, in maintaining maladaptive behaviour without seeing the bigger picture and therefore neglect the long-term consequences, or a focus on details in their eating behaviour. Also for these patients, the L-POST might provide an instrument to assess global-local biases.

Limitations and directions for further research

The current study has some limitations that can be addressed in follow-up research. First, test–retest reliability was evaluated with a relatively short time interval. Extending the time interval to assess generalizability over longer times would be advisable. This should take place in chronic patients in which spontaneous recovery is unlikely and who do not receive any treatment in between the two measures of the L-POST. Second, we only studied the internal structure of the L-POST in the healthy norm sample. When collecting more data of patients, one could evaluate correspondence of the factor structure in a patient sample and the healthy norm sample. Measurement invariance between groups can be assessed with confirmatory factor analyses. Third, when the factor structure has been established in both samples, calculating the factor scores for every patient and evaluating reliability of these scores might increase the interpretation of performance on the L-POST and enable more refined judgements about specific perceptual processes. Fourth, associations of L-POST scores with the lesion locations of the brain-damaged patients might provide us with more insights into the neural correlates of perceptual organization. Fifth, we did not include any of the process-specific tests on mid-level perception (e.g., 'Bender Gestalt Test', 'Street Completion Test') in our evaluation of convergent validity, as these only measure one aspect of what the L-POST aims to measure. However, it might be interesting to correlate performance on these tests with performance on the L-POST. Last, to apply the L-POST in other patient populations as suggested in the previous paragraph, the L-POST should be validated separately in these populations. In the longer run, we will develop additional, more psychophysical test procedures consisting of more items and possibly adaptive methods, which allow for specific follow-up testing, targeted at one or more subprocesses (e.g., based on the factor structure). In addition, we might develop a shortened version of the L-POST for which we will select the subtests with a high factor loading. However, this is only possible once we have confirmed a similar factor structure in a patient sample as in our healthy norm sample. We will also develop versions of the L-POST that are more suitable for children.

Conclusion

We can conclude that the L-POST is a reliable and valid instrument to screen for perceptual organization deficits in brain-damaged patients. In the healthy norm sample, Cronbach's alpha indicated adequate reliability, which was confirmed by high test–retest correlation in our patient sample. The strong embedding of the L-POST in the literature is reflected in the internal structure of the data. In addition, strong correlations with tests on mid-level perception and moderate correlations with tests of both mid- and high-level visual perception indicated good convergent validity. Furthermore, weak correlations with measures of spatial attention, executive functions, memory, language, number skills and praxis showed decent discriminant validity. The strong psychometric properties make the L-POST a valuable instrument for both clinicians and researchers interested in deficits in mid-level visual perception.

References

Anderson, P., Anderson, V., & Garth, J. (2001). Assessment and development of organizational ability: The Rey Complex Figure organizational strategy score (RCF-OSS). *The Clinical Neuropsychologist*, *15*, 81–94. doi:10.1076/clin.15.1.81.1905

Bach, M. (1996). The freiburg visual acuity test - automatic measurement of visual acuity. *Optometry and Vision Science*, *73*, 49–53. doi:10.1097/00006324-199601000-00008

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*, 600–609. doi:10.1162/089892903321662976

Barrett, G. V., Cabe, P. A., & Thornton, C. L. (1968). Visual functioning and embedded figures test performance. *Perceptual and Motor Skills*, *26*, 40. doi:10.2466/pms.1968.26.1.40

Bender, L. (1938). *A visual motor Gestalt test and its clinical use*. New York, NY: The American Orthopsychiatric Association.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–117. doi:10.1037/0033-295X.94.2.115

Binetti, G., Cappa, S. F., Magni, E., Padovani, A., Bianchetti, A., & Trabucchi, M. (1998). Visual and spatial perception in the early phase of Alzheimer's disease. *Neuropsychology*, *12*, 29–33. doi:10.1037/0894-4105.12.1.29

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex*, *16*, 183–191. doi:10.1093/cercor/bhi096

Bridge, H., Thomas, O. M., Minini, L., Cavina-Pratesi, C., Milner, A. D., & Parker, A. J. (2013). Structural and functional changes across the visual cortex of a patient with visual form agnosia. *The Journal of Neuroscience*, *33*, 12779–12791. doi:10.1523/JNEUROSCI.4853-12.2013

Coslett, H. B., & Saffran, E. (1991). Simultanagnosia. To see but not two see. *Brain*, *114*, 1523–1545. doi:10.1093/brain/114.4.1523

Dakin, S. C., & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron*, *48*, 497–507. doi:10.1016/j.neuron.2005.10.018

Della Salla, S., Laiacona, M., Trivelli, C., & Spinnler, H. (1995). Poppelreuter-Ghent's overlapping figures test: Its sensitivity to age, and its clinical use. *Archives of Clinical Neuropsychology*, *10*, 511–534. doi:10.1016/0887-6177(94)00049-V

Demeyere, N., Lestou, V., & Humphreys, G. W. (2010). Neuropsychological evidence for a dissociation in counting and subitizing. *Neurocase*, *16*, 219–237. doi:10.1080/13554790903405719

de-Wit, L. H., Kentridge, R. W., & Milner, A. D. (2009). Shape processing area LO and illusory contours. *Perception*, *38*, 1260–1263. doi:10.1068/p6388

de-Wit, L. H., Kubilius, J., Op de Beeck, H. P., & Wagemans, J. (2013). Configural Gestalts remain nothing more than the sum of their parts in visual agnosia. *i-Perception*, *4*, 493–497. doi:10.1068/i0613rep

Efron, R. (1969). What is perception? *Boston Studies in the Philosophy of Science*, *4*, 137–173. doi:10.1007/978-94-010-3378-7_4

Eliot, J., & Czarnolewski, M. (1999). A composite Gestalt completion test. *Perceptual and Motor Skills*, *89*, 294–300. doi:10.2466/pms.1999.89.1.294

Frith, U. (1989). *Autism: Explaining the enigma*. Oxford, UK: Blackwell.

Gilchrist, I. D., Humphreys, G. W., & Riddoch, M. J. (1996). Grouping and extinction: Evidence for low-level modulation of visual selection. *Cognitive Neuropsychology*, *13*, 1223–1249. doi:10.1080/026432996381737

Gilmore, G. C., Tobias, T. R., & Royer, F. L. (1985). Aging and similarity grouping in visual search. *Journal of Gerontology*, *40*, 586–592. doi:10.1093/geronj/40.5.586

Goodale, M. A., Jakobson, L. S., Milner, A. D., Perrett, D. I., Benson, P. J., & Hietanen, J. K. (1994). The nature and limits of orientation and pattern processing supporting visuomotor control in a visual form agnosic. *Journal of Cognitive Neuroscience*, *6*, 46–56. doi:10.1162/jocn.1994.6.1.46

Goodenough, D. R. (1976). The role of individual differences in field dependence as a factor in learning and memory. *Psychological Bulletin*, *83*, 675–694. doi:10.1037/0033-2909.83.4.675

Graetz, P., De Bleser, R., & Willmes, K. (1992). *Akense Afasie Test*. Lisse, The Netherlands: Swets & Zeitlinger.

Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences*, *20*, 106–111. doi:10.1016/S0166-2236(96)01002-8

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. doi:10.1080/10705519909540118

Humphrey, D. G., & Kramer, A. F. (1999). Age-related differences in perceptual organization and selective attention: Implications for display segmentation and recall performance. *Experimental Aging Research*, *25*, 1–26. doi:10.1080/036107399244110

Humphreys, G. W., Bickerton, W.-L., Samson, D., & Riddoch, M. J. (2012). *BCoS cognitive screen*. London, UK: Psychology Press.

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, *126*, 2463–2475. doi:10.1093/brain/awg248

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211. doi:10.3758/BF03212378

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*, 498–525. doi:10.2307/1418556

Koffka, K. (1922). Perception: An introduction to the Gestalt-Theorie. *Psychological Bulletin*, *19*, 531–585. doi:10.1037/h0072422

Kovács, I. (1996). Gestalten of today: Early processing of visual contours and surfaces. *Behavioural Brain Research*, *82*, 1–11. doi:10.1016/S0166-4328(97)81103-5

Kubovy, M., Holcombe, A. O., & Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, *35*, 71–98. doi:10.1006/cogp.1997.0673

Kubovy, M., & Wagemans, J. (1995). Grouping by proximity and multistability in dot lattices: A quantitative Gestalt theory. *Psychological Science*, *6*, 225–234. doi:10.1111/j.1467-9280.1995.tb00597.x

Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, *9*, 406–413.

Lawrence, A. D., Watkins, L. H. A., Sahakian, B. J., Hodges, J. R., & Robbins, T. W. (2000). Visual object and visuospatial cognition in Huntington's disease: Implications for information processing in corticostriatal circuits. *Brain*, *123*, 1349–1364. doi:10.1093/brain/123.7.1349

Lopez, M. N., Charter, R. A., Oh, S., Lazar, M. D., & Imperio, S. M. (2005). Psychometric properties of the benton visual form discrimination test. *Applied Neuropsychology*, *12*, 19–23. doi:10.1207/s15324826an1201_4

Lopez, C., Tchanturia, K., Stahl, D., Booth, R., Holliday, J., & Treasure, J. (2008). An examination of the concept of central coherence in women with anorexia nervosa. *The International Journal of Eating Disorders*, *41*, 143–152. doi:10.1002/eat.20478

Lopez, C. A., Tchanturia, K., Stahl, D., & Treasure, J. (2008a). Central coherence in women with bulimia nervosa. *The International Journal of Eating Disorders*, *41*, 340–347. doi:10.1002/eat.20511

Lopez, C. A., Tchanturia, K., Stahl, D., & Treasure, J. (2008b). Central coherence in eating disorders: A systematic review. *Psychological Medicine*, *38*, 1393–1404. doi:10.1017/S0033291708003486

Machilsen, B., & Wagemans, J. (2011). Integration of contour and surface information in shape detection. *Vision Research*, *51*, 179–186. doi:10.1016/j.visres.2010.11.005

Mendez, M. F., Mendez, M. A., Martin, R., Smyth, K. A., & Whitehouse, P. J. (1990). Complex visual disturbances in Alzheimer's disease. *Neurology*, *40*(3 Part 1), 439. doi:10.1212/WNL.40.3_Part_1.439

Meyers, J. E., & Meyers, K. R. (1996). *Rey complex figure test and recognition trial*. Lutz, FL: Psychological Assessment Resources.

Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., . . . Terazzi, E. (1991). Perception and action in "visual form agnosia". *Brain*, *114*, 405–428. doi:10.1093/brain/114.1.405

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, *9*, 429–434. doi:10.1038/nn1641

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, *18*, 55–68. doi:10.1068/p180055

Norris, G., & Tate, R. L. (2000). The behavioural assessment of the dysexecutive syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation*, *10*, 33–45. doi:10.1080/096020100389282

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. (3rd ed.). New York, NY: McGraw-Hill.

Nys, G. M. S., van Zandvoort, M. J. E., de Kort, P. L. M., van der Worp, H. B., Jansen, B. P. W., Algra, A., . . . & Kappelle, L. J. (2005). The prognostic value of domain-specific cognitive abilities in acute first-ever stroke. *Neurology*, *64*, 821–827. doi:10.1212/01.WNL.0000152984.28420.5A

Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *The Journal of Neuroscience*, *26*, 13025–13036. doi:10.1523/JNEUROSCI.2481-06.2006

Op de Beeck, H. P., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *The Journal of Neuroscience*, *28*, 10111–10123. doi:10.1523/JNEUROSCI.2511-08.2008

Pisella, L., André, V., Gavault, E., Le Flem, A., Luc-Pupat, E., Glissoux, C., . . . Gonzalez-Monge, S. (2012). A test revealing the slow acquisition and the dorsal stream substrate of visuo-spatial perception. *Neuropsychologia*, *51*, 106–113. doi:10.1016/j.neuropsychologia.2012.11.015

R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical computing. Retrieved from www.R-project.org

Riddoch, M. J., & Humphreys, G. W. (1987). A case of integrative visual agnosia. *Brain*, *110*, 1431–1462. doi:10.1093/brain/110.6.1431

Riddoch, M. J., & Humphreys, G. W. (1993). *Birmingham object recognition battery*. London, UK: Psychology Press.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience*, *29*, 203–227. doi:10.1146/annurev.neuro.29.051605.112939

Rosseel, Y. (2012). lavaan: An R package for structural equation. *Journal of Statistical Software*, *48*, 1–36.

Schenkel, L. S., Spaulding, W. D., & Silverstein, S. M. (2005). Poor premorbid social functioning and theory of mind deficit in schizophrenia: Evidence of reduced context processing? *Journal of Psychiatric Research*, *39*, 499–508. doi:10.1016/j.jpsychires.2005.01.001

Segaert, K., Nygård, G. E., & Wagemans, J. (2009). Identification of everyday objects on the basis of kinetic contours. *Vision Research*, *49*, 417–428. doi:10.1016/j.visres.2008.11.012

Silverstein, S. M., & Keane, B. P. (2011). Perceptual organization impairment in schizophrenia and associated brain mechanisms: Review of research from 2005 to 2010. *Schizophrenia Bulletin*, *37*, 690–699. doi:10.1093/schbul/sbr052

Southgate, L., Tchanturia, K., & Treasure, J. (2005). Building a model of the aetiology of eating disorders by translating experimental neuroscience into clinical practice. *Journal of Mental Health*, *14*, 553–566. doi:10.1080/09638230500347541

Spillmann, L. (1999). From elements to perception: Local and global processing in visual neurons. *Perception*, *28*, 1461–1492. doi:10.1068/p2763

Tippett, L. J., Blackwood, K., & Farah, M. J. (2003). Visual object and face processing in mild-to-moderate Alzheimer's disease: From segmentation to imagination. *Neuropsychologia*, *41*, 453–468. doi:10.1016/S0028-3932(02)00140-9

Torfs, K., Vancleef, K., Lafosse, C., Wagemans, J., & de-Wit, L. H. (2014). The Leuven Perceptual Organization Screening Test (L-POST), an online test to assess mid-level visual perception. *Behavior Research Methods*, *46*, 472–487. doi:10.3758/s13428-013-0382-6

Trick, L. M. (2008). More than superstition: Differential effects of featural heterogeneity and change on subitizing and counting. *Perception & Psychophysics*, *70*, 743–760. doi:10.3758/PP.70.5.743

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*, 371–387. doi:10:1167/2.5.2

Uhlhaas, P. J., Phillips, W. A., Mitchell, G., & Silverstein, S. M. (2006). Perceptual grouping in disorganized schizophrenia. *Psychiatry Research*, *145*, 105–117. doi:10.1016/j.psychres.2005.10.016

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, *138*, 1172–1217. doi:10.1037/a0029333

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*, 1218–1252. doi:10.1037/a0029334

Wagemans, J., Wichmann, F. A., & Op de Beeck, H. P. (2005). Visual perception I: Basic principles. In K. Lamberts & R. Goldstone (Eds.), *Handbook of cognition* (pp. 3–47). London, UK: Sage.

Warrington, E. K., & James, M. (1986). Visual object recognition in patients with right-hemisphere lesions: Axes or features? *Perception*, *15*, 355–366. doi:10.1068/p150355

Warrington, E. K., & James, M. (1991). *The visual object and space perception battery*. Bury St. Edmunds, UK: Thames Valley Test Company.

Warrington, E. K., & Taylor, A. M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex*, *9*, 152–164. doi:10.1016/S0010-9452(73)80024-3

Weiskrantz, L. (1990). *Blindsight*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780198521921.001.0001

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 71–88). London, UK: Kegan Paul, Trench, Trubner & Co.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, *24*, 55–62. doi:10.1016/0042-6989(84)90144-5

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome*. Bury St Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Cockburn, J., & Baddeley, A. (1989). *The Rivermead behavioural memory test*. Bury St Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Cockburn, J., Baddeley, A., & Hiorns, R. (1991). *The Rivermead behavioural memory test*: *Supplement 2*. Bury St Edmunds, UK: Thames Valley Test Company.

Witkin, H. A. (1962). *Psychological differentiation: Studies of development*. New York: Wiley.

Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain*, *106*, 313–340. doi:10.1093/brain/106.2.313

Received 3 April 2014; revised version received 22 May 2014

Supporting Information

The following supporting information may be found in the online edition of the article:

Appendix S1. Descriptions of the patients.
Appendix S2. Average performance and variance–covariance matrix of the healthy norm sample.
Appendix S3. Correlations between performance on the BORB and VOSP subtests and visual extinction (BCoS), executive functions (BADS), and language (AAT).