Neuropsychologia 46 (2008) 2463–2475

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Verbal numerosity estimation deficit in the context of spared semantic representation of numbers: A neuropsychological study of a patient with frontal lesions

Susannah K. Revkin[a,b,c,*], Manuela Piazza[a,b,c,d], Véronique Izard[a,b,c,e], Laura Zamarian[f], Elfriede Karner[f], Margarete Delazer[f]

[a] INSERM, U562, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France
[b] CEA, DSV/I2BM, NeuroSpin Center, F-91191 Gif/Yvette, France
[c] Univ Paris-Sud, IFR49, F-91191 Gif/Yvette, France
[d] Functional NeuroImaging Laboratory, Center for Mind Brain Sciences, University of Trento, Italy
[e] Department of Psychology, Harvard University, Cambridge, USA
[f] Innsbruck Medical University, Austria

ARTICLE INFO

Article history:
Received 30 August 2007
Received in revised form 12 April 2008
Accepted 16 April 2008
Available online 22 April 2008

Keywords:
Frontal
Quantity
Approximate
Visual
Executive processes

ABSTRACT

Patients with frontal lobe damage have been shown to produce implausible answers in cognitive estimation, a task requiring approximate answers to quantity-related questions of general semantic knowledge. We investigated a patient with frontal lobe damage who presented executive deficits and difficulties in cognitive estimation. The patient also showed difficulties in verbal numerosity estimation (approximately evaluating the quantity of visually presented sets of items), as he produced extreme answers well outside healthy participants' range of answers. A series of tasks evidenced intact number processing and well preserved semantic representation of numbers. Detailed investigation of estimation processes suggested a deficit at the level of translation from an intact semantic representation of numbers to output, whether verbal or non-symbolic. This case study allows disentangling different processes involved in estimation and contributes to a better understanding of the cognitive estimation deficits frequently reported for patients with frontal lesions.

1. Introduction

It has long been known that focal frontal lobe damage can sometimes cause relatively isolated cognitive deficits, which almost go unnoticed, as general intellectual capacities can be spared. One striking finding revealed that some patients with frontal lobe damage, whose general intellectual abilities were intact, presented specific difficulties in cognitive estimation, the capacity to give approximate answers to questions of general semantic knowledge for which no precise answer is readily known (Shallice & Evans, 1978). Indeed, these patients' performance, when presented with questions pertaining for example to the size, height, or weight of objects, was characterized by extremely implausible answers (example of an answer in response to the question "what is the length of an average man's spine?": "between 4 and 5 feet"). As intellectual capacities were spared, this type of deficit was interpreted as resulting from impairment of selective and regulative processes attributed to the frontal lobes (selecting possible answers, checking for the plausibility of each answer, etc.), rather than from degradation of general semantic knowledge.

On the other hand, patient studies (Brand, Kalbe, & Kessler, 2002a; Della Sala, MacPherson, Phillips, Sacco, & Spinnler, 2004; Mendez, Doss, & Cherrier, 1998; Taylor & O'Carroll, 1995, Experiment 3) have brought evidence that cognitive estimation deficits may not be specific to patients with focal frontal lobe damage. Indeed, cognitive estimation can also be impaired in patients with posterior lesions: in these cases the cognitive estimation deficit supposedly reflects impairment of general knowledge itself (semantic memory) known to be mainly sub-served by the temporal lobe.

Frontal lobe patients are known to show deficits in cognitive estimation, but since this task clearly also requires semantic knowledge (knowledge of the world, such as distances, weights or lengths) it is often difficult to disentangle the contribution of

* Corresponding author at: Inserm-CEA Cognitive Neuroimaging Unit, CEA/SAC/DSV/I2BM/NeuroSpin, Bât 145, Point Courrier 156, F-91191 Gif/Yvette, France. Tel.: +33 1 69 08 83 90; fax: +33 1 69 08 79 73.
E-mail address: susannahrevkin@gmail.com (S.K. Revkin).

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.04.011

impaired semantic knowledge and impaired estimation processes. *Verbal numerosity estimation*, that is, giving a verbal estimate of the quantity of a set of items, differs from cognitive estimation because it requires a *perceptual* judgment of quantity (as would, for example, judging the length of the experimenter's spine).

Not many studies have specifically investigated the cerebral bases of verbal numerosity estimation. Three neuropsychological studies (Delazer, Karner, Zamarian, Donnemiller, & Benke, 2006; Pesenti, Thioux, Samson, Bruyer, & Seron, 2000; Warrington & James, 1967) suggest a role of parietal structures, in particular the right parietal lobe, in verbal numerosity estimation. This makes sense, as mere perception of or perceptual comparative judgments of numerosity (without a verbal output; e.g., comparing the quantity of two sets of dots) have been linked to parietal structures through imaging studies (e.g. Cantlon, Brannon, Carter, & Pelphrey, 2006; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Piazza, Pinel, Le Bihan, & Dehaene, 2007). Unlike cognitive estimation, verbal numerosity estimation does not rely on general semantic knowledge (sub-served by temporal structures), but rather on the intact processing of domain specific numerical representations, that is representation of numerosity (semantic representation of numbers) sub-served by parietal structures, as these studies suggest (for a review on numerical processing and the parietal lobes, see Dehaene, Piazza, Pinel, & Cohen, 2003). Thus, investigating verbal numerosity estimation in a patient with perfect processing of numerosities would allow specifically studying estimation processes without confounding deficits in semantic knowledge or number representation.

Could verbal numerosity estimation be impaired following focal frontal lobe damage, as is the case for cognitive estimation? To our knowledge, this question has not been specifically studied in controlled conditions. Similarly to the results of studies pertaining to cognitive estimation (Della Sala et al., 2004; Shallice & Evans, 1978; Smith & Milner, 1984, Experiment 1), one could expect impairments in verbal numerosity estimation in patients with frontal lobe damage, as it also represents a task in which no exact answer is readily available (in contrast to counting), and calls upon the selection of a response among a theoretically infinite range of possibilities. Here we report a study in which we tested this hypothesis by administering a verbal numerosity estimation test to a patient with focal frontal lesions, a cognitive estimation deficit and executive impairments.

If a verbal numerosity estimation deficit should arise, it would be of importance to determine whether it is linked to impairment at the level of the semantic representation of numbers. Although the semantic representation of numbers has been linked to parietal structures as discussed above, some studies suggest a possible additional involvement of the fronto-lateral cortex (in monkeys: Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2004, for a review, see Nieder, 2005; in human healthy adults: e.g. Piazza, Mechelli, Price, & Butterworth, 2006). However, we did not aim to test this hypothesis, and, importantly, the patient investigated in this study presented frontal lesions which did not extend to the fronto-lateral cortex. Thus, we hypothesized that the semantic representation of numbers should be spared in such a patient. We tested this by administering tasks known to recruit the semantic representation of numbers, and which do not require a verbal output: dots comparison, dots addition and comparison, digit comparison, and number-size Stroop digit comparison.

Another level which should be investigated, should a verbal numerosity estimation deficit arise, is external calibration. Calibration characterizes the spontaneous mapping from the approximate semantic representation of numbers to a verbal response grid during verbal numerosity estimation. Healthy subjects have been shown to be poorly calibrated, that is, they present coherent estimates (estimates which increase as numerosity increases) but systematically under- or overestimate the presented numerosities (Izard & Dehaene, 2008; Minturn & Reese, 1951). External calibration (showing an example of a numerosity concurrently to the correct verbal response) has been shown to improve estimates such that under- or overestimation is significantly reduced (Izard & Dehaene, 2008; Minturn & Reese, 1951). It was suggested that this external calibration process was probably a mix between strategic and automatic adjustment of verbal responses to the semantic representation of numbers (Izard & Dehaene, 2008). Would a patient presenting a verbal numerosity estimation deficit benefit from external calibration, similarly to healthy subjects? We wished to also address this question and therefore also tested verbal numerosity estimation with external calibration. We hypothesized that a patient with frontal lesions and executive deficits would improve less following external calibration, considering that this process might involve a strategic (executive) component.

Finally, it would be important to rule out the possibility that a verbal numerosity estimation deficit might occur in relation to a specific impairment at the verbal output level. If the deficit persists with different output modalities, and if the semantic representation of numbers is intact, this would suggest that the impairment is situated at the level of translation from representation to output. We investigated this first by testing numerosity estimation with a forced-choice paradigm presenting symbolic output other than number words (forced-choice estimation "from dots to digits"); and secondly, with a forced-choice paradigm presenting non-symbolic output (forced-choice estimation "from digits to dots"). We hypothesized that the estimation deficit should extend to these other outputs, reasoning that the impairment should be situated at the level of translation from semantic representation to output, as cognitive estimation deficits following frontal lobe damage seem linked to impairments at this level (selective and regulative deficits).

2. Methods and results

2.1. Case description

The patient we examined was a 28-year old right-handed native German speaking man who had completed polytechnic studies and trained as an engines fitter. He was the beneficiary of an incapacity pension, following a car accident about 8 years prior to testing, that had caused left frontal substance defect. About 2 years prior to testing, the patient had suffered a second accident (a fall down some stairs), causing right cerebral contusions. A computed tomography (CT) scan taken during the testing period showed left fronto-polar to fronto-basal damage (see Fig. 1).

Because of the recent occurrence of epileptic Grand Mal seizures, he underwent routine neuropsychological testing. Following the testing, he was asked if he wished to participate in this study. The patient agreed and gave his informed written consent prior to his inclusion in the study.

2.2. Healthy participants

An initial group of 15 healthy unpaid volunteers (five men) was used as a comparison of the patient's results on most tasks. The volunteers were 21–43 years of age (mean age = 26.87 years). For one task (see Section 2.9), data were collected from a second group of 15 healthy participants (10 men), 5 of which were participants from the first group. Participants of this second group were 24–37 years of age (mean = 28.00 years). Participants of both groups were all native German speakers. Finally, for one other task (see Section 2.6.1) we used control data collected from 18 healthy French-speaking paid volunteers (8 men; mean age = 24.94 years, ranging from 18 to 38), participating in another study. We used these data even though they had been collected from French speakers, because the task did not call for verbal responses. Participants from all three groups were right-handed and of similar educational level (all university students or graduates), and gave their informed consent prior to their inclusion in the study.

2.3. Neuropsychological examination

A neuropsychological evaluation of the patient was carried out two days before numerical testing began (all results are reported in Table 1).

The patient presented a slight deficit in verbal long-term memory (learning and recall difficulties, consolidation and recognition being intact), in verbal production

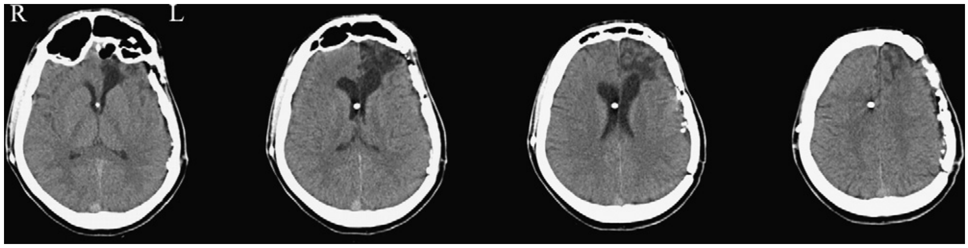

Fig. 1. CT-scan showing left fronto-basal to fronto-polar damage.

(categorical, phonological and alternating phonological fluency tests), and a deficit in decision making (IOWA gambling task). The patient also presented inhibition difficulties in a go-no-go task, attention fluctuations in a phasic alertness test, and slow (although sufficiently accurate) performance in different tasks (divided attention test; complex mental calculation test; copy of a complex geometrical figure). He also presented extremely positive scores on the novelty component of a sensation-seeking scale (although scoring within normal range on the intensity component), and occasional behavior which was contextually inadequate or impulsive. Figural long-term memory was globally spared, free recall being within normal range at a long delay although impaired at a short delay, and recognition being intact. Planning and problem solving was generally spared as the patient's number of solved trials was within normal range, although he made an excessive number of errors in this test. There was no deficit in categorical alternated verbal fluency, verbal span and working memory (digit spans forward and backward), alertness without warning, cognitive flexibility, and in all subtests of a short battery investigating executive functions (FAB). Finally, verbal IQ was estimated at 91, a score that was in the normal range. In sum, the patient presented executive impairments compatible with and typical of focal frontal lobe damage. The experimental testing reported in the next section was carried out over four sessions which covered a period of 2 months. All computerized tasks were programmed and administered using e-prime software (Schneider, Eschman, & Zuccolotto, 2002).

2.4. *Cognitive estimation*

The "Test zum kognitiven Schätzen" was administered (TKS, Brand, Kalbe, & Kessler, 2002b) and showed a marked impairment (6 correct/16), visible in all four categories (size: 2/4; weight: 1/4; numerosity: 1/4; time: 2/4), half the time due to underestimation and the other half to overestimation. For example, when shown a picture of a pair of glasses and asked to estimate its weight, the patient replied "2 grams" (acceptable range = 24–130 g). Or, when shown a picture of several flowers, and asked how many there were, he gave the answer "50–60", a response well above the acceptable range (15–31). In most sub-sections of the TKS, estimation is prompted either by pictures of objects which, importantly, are not represented in real size (sub-sections size and weight), or by sentences (time). In these sub-sections, subjects therefore must rely mainly on cognitive as opposed to perceptual processes (thus the term "cognitive estimation"). In contrast, the numerosity sub-section of the TKS does call upon perceptual processes, and can in fact be considered a test of verbal numerosity estimation. However, this sub-section contains only four items and does not control for confounds with non-numerical variables, which must be taken into account to insure that estimation is based on numerosity in these trials. Also, although the patient's results in this sub-test suggest impaired estimation of numerical quantity, they are not sufficient to situate at what level the deficit occurs (semantic representation of numbers, or translation from semantic representation of numbers to output, etc.). We further investigated this with the help of a set of different numerical tasks, after first re-testing verbal numerosity estimation in a more controlled task with more items.

2.5. *Verbal numerosity estimation without external calibration*

2.5.1. *Method*

The patient was presented with sets of dots which represented the following 11 numerosities: 10, 13, 17, 22, 29, 37, 48, 63, 82, 106, and 138. He was instructed to estimate as accurately as possible the quantity of dots present in the display without counting. In order to prevent him from using non-numerical parameters that usually co-vary with numerosity (e.g. density of the dots or the size of the area of the envelope of the cloud of dots), density was held constant across numerosities for half of the stimuli, and the area of the envelope of the cloud of dots was held constant for the other half of the stimuli (with randomization of this variable of control of non-numerical parameters across trials).[1] The test was administered in one session of three blocks resulting in a total amount of 66 trials (6 per numerosity, in random order). The dots were black (∼0.2°) and appeared for 700 ms in a white disc (diameter of 8.4°) which remained on the screen throughout the experiment. The patient entered his response using the computer keyboard. After each response, 1400 ms elapsed before the next set of dots.

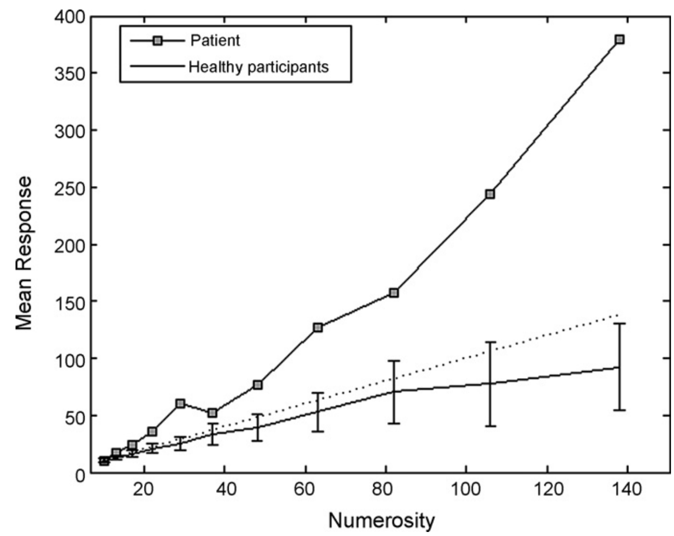

Fig. 2. Patient's vs. healthy participants' performance in verbal numerosity estimation without external calibration. (Error bars represent ±1 standard deviation; the dashed line represents theoretical exact correspondence between numerosity and mean response; mean responses can therefore be visualized as under- or overestimates if they fall respectively below or above the dashed line.)

We analyzed responses to detect extreme answers (under- or overestimation). Responses and their variability were also used to detect signatures of estimation processes. One would typically expect a correlation between the presented numerosities and the patient's mean responses. One would further expect "scalar variability", which is a classical signature of estimation processes (Gallistel & Gelman, 1992; Izard & Dehaene, 2008; Whalen, Gallistel, & Gelman, 1999), and which characterizes the fact that estimation judgments become less precise as numerosity increases in such a way that the variability in responses increases *proportionally* to the increase in mean response. This results in a stable variation coefficient (standard deviation of responses divided by mean response) across numerosities. This measure represents a quantitative estimate of the precision of numerical estimation, so we also tested if the patient's value was higher than those of healthy participants, which would indicate a reduced precision of numerical estimates. Finally, response times were also analyzed as they are typically found to be flat across numerosities during estimation. In sum, we used different measures allowing us to detect both qualitative and quantitative differences between the patient and healthy participants.

2.5.2. *Patient's results*[2]

As concerns extreme answers, the patient's responses (M = 88.19, S.D. = 74.81; see Fig. 2), were consistently larger than the correct response across numerosities (and ranged from 9 to 500, or even 700 for numerosity 138), reflecting a clear tendency to overestimate. However, it is important to note that further analyses

[1] For this task, as well as the tasks described in Sections 2.6.2, 2.7, 2.8, and Section 2.9, analyses pertaining to non-numerical parameters are reported in Section 2.11. "Sensitivity to non-numerical parameters".

[2] Unless specified otherwise, we report results and analyses excluding data from the extremes numerosities (10 and 138) to avoid noise from anchoring effects.

Table 1
Neuropsychological background tests results

	Patient	Normal range
Verbal intelligence		
Premorbid IQ (Lehrl, Merz, Burkard, & Fischer, 1991)	91	81–145[a]
Memory		
Verbal memory (VLMT-A; Helmstaedter & Durwen, 1990)		
Verbal learning (total)[b]	38	43–75[a]
Free recall, short delay[b]	7	9–15[a]
Free recall, long delay[b]	7	9–15[a]
Loss after temporal delay (consolidation)	3	−1 to 4[a]
Recognition	15	13–15[a]
Figural memory (RCFT; Rey, 1941; Spreen & Strauss, 1998)		
Free recall, short delay[c]	14.5	18.5–36[a]
Free recall, long delay	21	18.5–36[a]
Recognition	21	20–24[a]
Attention		
Digit span forward (WMS-R; Wechsler, 1987)	6	6–11[a]
Alertness (TAP; Zimmermann & Fimm, 2002)		
Alertness without warning (median, ms)	227	<306.98[a]
(S.D., ms)	33	<72.33[a]
Alertness with warning (median, ms)	199	<285.49[a]
(S.D., ms)[c]	138	<76.45[a]
Phasic alertness (score)	0.13	−0.07 to 0.42[a]
Divided attention (TAP)		
Median (ms)[b]	769	<628.35[a]
S.D. (ms)[c]	385	<137.96[a]
Errors	0	0–5[a]
Executive functions		
Digit span backward (WMS-R, Wechsler, 1987)	5	5–10[a]
Complex mental calculation (GDAE; Jackson & Warrington, 1986) (scaled score)	8	7–17[a]
Verbal fluency (RWT; Aschenbrenner, Tucha, & Lange, 2001)		
Categorical verbal fluency (animals/min)[b]	15	16–35[a]
Phonological verbal fluency (s-words/min)[b]	8	9–24[a]
Alternated verbal fluency (alternation sports-fruits/min)	12	11–22[a]
Alternated verbal fluency (alternation h-words vs. t-words/min)[b]	6	10–21[a]
Planning and problem-solving (TOL, German version; Kohler & Beck, 2004; Kohler, Beck, & Hohnecker, 2003) (trials)		
Solved trials	6	4–6[a]
Errors[c]	7	0–1[a]
Cognitive flexibility (OMO; Flowers & Robertson, 1985) (errors)	0	0–2[a]
Inhibitory control (go-no-go task, computerized version; adapted from Fox, Michie, Wynne, & Maybery, 2000)		
Go correct (%)	98.9	98.9–100[a]
No Go correct (%)[c]	63.3	78.7–97.0[a]
FAB (Dubois, Slachevsky, Litvan, & Pillon, 2000) (total score)	17	16.1–18[a]
Conceptualization	3	
Mental flexibility	3	
Motor programming	3	
Sensitivity to interference	3	
Inhibitory control	2	
Environmental flexibility	3	
IOWA gambling task (Bechara, Damasio, Damasio, & Anderson, 1994; Bechara, Tranel, & Damasio, 2000)		
Block 1–5 (total draws from favorable decks)[c]	34	67.6/S.D. 9.9[d]
Constructive abilities		
Copying a complex geometrical figure (RCFT; Rey, 1941; Spreen & Strauss, 1998)		
Score	34	33.5–36[a]
Duration (s)[c]	487	<289[a]
Sensation seeking		
Intensity (AISS-D; Roth, Schumacher, & Arnett, 2003)	28	28.31/S.D. 0.96[d]
Novelty (AISS-D)[b]	26	24.42/S.D. 0.98[d]

[a] Within 1.5S.D. from the mean.
[b] Patient's result below/above 1.5S.D. from the mean of standardized norms.
[c] Patient's result below/above 2S.D. from the mean of standardized norms.
[d] Mean and standard deviation of healthy individuals.

revealed the presence of signatures of estimation processes. Indeed, responses correlated positively with numerosity ($r = 0.74$, $p < 0.01$), indicating that they were not random. Also, inspection of the response distribution indicated that the variability in responses tended to increase as numerosity increased, and analyses revealed that the patient's mean variation coefficient ($M = 0.43$, S.D. = 0.19) was constant across numerosities ($R = 2.78$, $p = 0.12$; intercept = 0.25, slope = 0.002), confirming the presence of scalar variability. Moreover, response times (RTs), computed after having removed outliers (which were defined as RTs above or below two standard devi-

Table 2
Comparison of patient's vs. healthy participants' results in verbal numerosity estimation without external calibration

Response	Patient	Healthy participants		t-value (d.f. = 14)	p-value (two-tailed)
		Mean	S.D.		
Mean**	88.19	38.47	11.91	4.04	<0.01
Constant area*	62.11	35.50	10.81	2.38	<0.05
Constant density**	114.26	37.60	11.32	6.56	<0.01
Numerosity-response correlation coefficient*	0.74	0.88	0.04	−2.33	<0.05
Constant area	0.73	0.87	0.06	−1.70	0.11
Constant density	0.88	0.91	0.04	−0.70	0.50
Mean variation coefficient**	0.43	0.21	0.05	4.26	<0.01
Constant area*	0.33	0.20	0.05	2.52	<0.05
Constant density**	0.38	0.19	0.05	3.68	<0.01
RT (ms)					
Overall*	5661	3509	888	2.35	<0.05

* Patient significantly differs from healthy participants' at $p < 0.05$.
** At $p < 0.01$.

ations of the mean[3]; $M = 5661$ ms, S.D. = 1928 ms), were stable across numerosities ($R < 0.001$, $p = 0.99$; intercept = 5658, slope < 1), as expected in an estimation process.

2.5.3. Comparison to healthy participants

Although the patient presented signatures of estimation processes, there were quantitative differences with healthy participants. First of all, the patient's overestimation can be considered extreme, as his mean response was statistically higher than healthy participants' (see Table 2 for all results of this section); the patient's mean response to each numerosity exceeded healthy participants' by over 2 standard deviations (see Fig. 2). Also, the patient's numerosity-response correlation was significantly lower than healthy participants'. His mean variation coefficient across numerosities was statistically higher than healthy participants'. Finally, his mean RT was significantly slower than healthy participants'.

In sum, the patient's performance presented classical signatures of estimation processes: his responses correlated with numerosity and respected scalar variability (stable variation coefficient across numerosities), and his RTs were flat across numerosities. However, his results differed quantitatively from healthy participants', showing in particular larger responses (a clear pattern of overestimation), as well as a larger variation coefficient, indicative of less precise estimation.

In order to better understand the origin of the patient's estimation deficit, we administered several tests tapping into the semantic representation of numbers, in order to rule out a deficit at this level. We hypothesized that this level would be intact, as the patient's lesions did not involve parietal structures which are known to be implicated in the semantic representation of numbers.

2.6. Semantic representation of numbers

We administered several tasks tapping into the semantic representation of numbers but not requiring the production or selection of a symbolic output, in order to determine whether the estimation deficit was due to a core numerical deficit.

2.6.1. Dots comparison

2.6.1.1. Method. The patient was presented with two clouds of dots, and was asked to judge as accurately and as fast as possible which one contained the most dots. On each trial, one cloud contained a fixed numerosity (16 for half the trials, 32 for the other half) and the other cloud ("varying numerosity") contained a numerosity which was smaller or larger than the fixed numerosity by one of four possible ratios: 1.06, 1.13, 1.24, or 1.33. Thus, comparison difficulty was manipulated (the smaller the ratio, the harder the comparison). These variables were randomized across blocks. The patient responded by pressing the mouse button on the same side as the larger cloud (using his left or right index). The black dots (visual angle $\sim 0.25^{\circ}$) were present on the screen until the patient responded, and appeared after a delay of 1400 ms in two white discs (disc diameter = 7.2°; distance between the discs = 1.8°) on a black background on either side of a central white fixation spot (0.2°). On half the trials, the dot size of the varying numerosity cloud remained constant, and on the other half, the size of the area occupied by the varying numerosity cloud remained constant, whereas the fixed numerosity clouds varied on both parameters at once. This was designed to prevent the patient from basing his performance on these non-numerical parameters. First 16 training trials were performed with accuracy feedback. The patient performed a total of 128 trials over four blocks in one session (32 trials per ratio category). Accuracy was measured and analyzed in relation to ratio. Indeed, performance in non-symbolic numerical comparison typically improves with the ratio of the numbers to be compared (distance effect). Accuracy was also used to calculate the estimate of the internal Weber Fraction (w), a measure of the precision of underlying numerical representation. We tested if the patient's w value was higher than that of healthy participants, which would indicate reduced discrimination precision.

2.6.1.2. Patient's results. Overall accuracy was good (87% correct). Data from both fixed numerosities were collapsed, and analyzed in relation to the ratio of the varying numerosities. Results (see Fig. 3A) showed a distance effect, as expected: accuracy was lower for pairs where the varying numerosity differed from the fixed numerosity by a smaller ratio, and gradually increased as ratio increased ($R = 14.17$, $p < 0.01$; intercept = −0.58; slope = 1.23). This was also apparent as a correlation between ratio and accuracy ($r = 0.84$, $p < 0.01$). Accuracy scores were also used to calculate w, using a method previously described (maximum likelihood decision model, Supplemental Data from Piazza et al., 2004). The subject's w was of 0.14.

2.6.1.3. Comparison to healthy participants[4]. For each measure of this test, whether qualitative (presence of a distance effect) or quantitative (overall accuracy, w), the patient was found to be statistically comparable to healthy participants (see Table 3 for all results of this section, and also Fig. 3A). Overall accuracy in the healthy participants' group was slightly lower than the patient's although this difference was not significant. The patient's distance effect, as measured by the slope of the regression of accuracy against ratio, and as the correlation between ratio and accuracy, was not significantly different from healthy participants'.[5] The patient's w was slightly lower than healthy participants, indicating a slightly higher discrimination precision, although this difference was not significant.

2.6.2. Dots addition and comparison

2.6.2.1. Method. In each trial, the patient was presented with three large sets of dots one after the other, the first two being yellow and the third blue. He was required to mentally "add" the two yellow sets and compare this result to the blue set, in order to determine whether there were altogether more yellow dots or more blue dots (each yellow set could contain 10–53 dots; each blue set, 30–80 dots). He was asked not to count, but to estimate as accurately and as quickly as possible the number of dots in each set and respond by pressing the left mouse button with his left index for a larger quantity of yellow dots, and the right mouse button with his right index for a larger quantity of blue dots. The ratio between the two numerosities that constituted each comparison pair (i.e. between the result of the addition of the yellow sets, and the blue set) was manipulated to form three ratio categories, from which stimuli were selected randomly across trials: ∼1.3, ∼1.5, 2.0. Each session began with 10 training trials with feedback. The background was black and stayed empty (600 ms) before each set of dots appeared centrally (400 ms). If the patient did not respond during the presentation of the last cloud of dots, a black screen appeared and remained until he did respond. Half the sets were of constant density and dot size (mean visual angle of each dot = 0.2°), and the other half of constant total occupied area (area of

[3] Outlier RTs were defined in this way for all subsequent analyses of this study.

[4] Healthy participants performed the task in the same conditions as the patient except that they performed twice as many trials over two sessions.

[5] Following Crawford and Garthwaite (2004), we wished to statistically compare the slope of the patient's regression to that of healthy participants'. Given that there were differences among the healthy participants' error variances, this test was not applicable and we instead determined whether the patient's slope was within 2 S.D.s of healthy participants' slopes. However, we also computed correlations as a measure of the distance effect, to statistically compare the patient's measure with healthy participants' (see Crawford, Garthwaite, Howell, & Venneri, 2003).

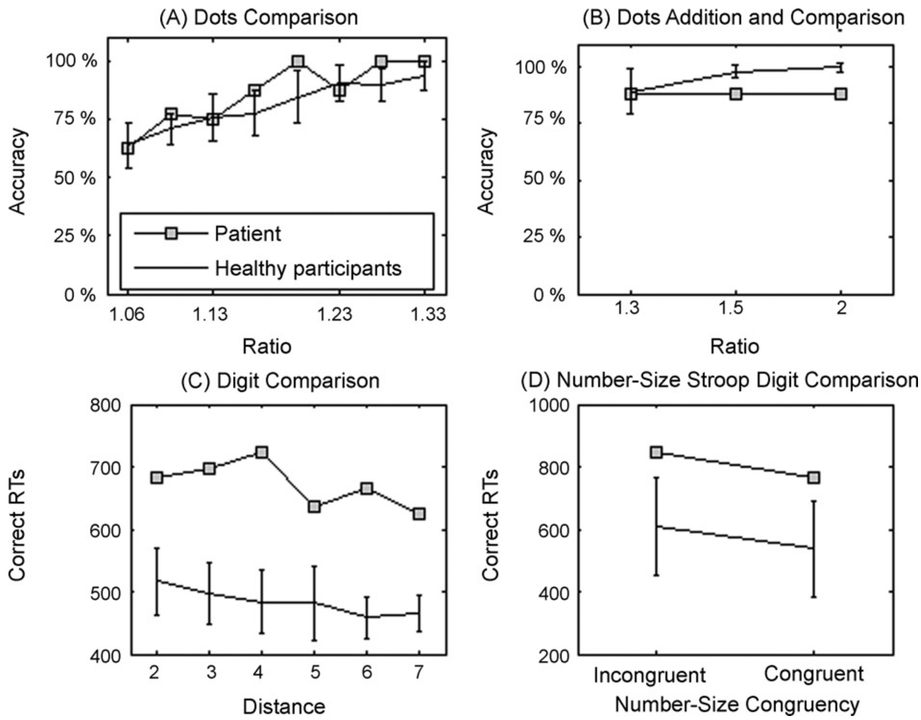

Fig. 3. Comparison of patient's vs. healthy participants' performance in the four tasks tapping into the semantic representation of numbers. (A) Dots comparison: accuracy by ratio; (B) dots addition and comparison: accuracy by ratio; (C) digit comparison: distance effect on correct response times (RTs); (D) number-size Stroop digit comparison: number-size incongruence effect on correct RTs.

about 5.7°; randomization of this variable across trials), to control for these non-numerical parameters. Data were gathered in one session of 48 trials, amounting to 16 presentations per ratio category. In half the trials the yellow quantity was larger than the blue quantity (randomization across trials). Accuracy and RTs were measured. Accuracy was also analyzed in relation to ratio, to detect a distance effect, as in the previous task.

2.6.2.2. Patient's results. Overall accuracy was of 88%. There was no distance effect, as accuracy did not vary at all in the different ratio conditions (difference ratio 2 minus ratio 1.3 = 0) (see Fig. 3B).[6] Correct RTs (computed after having removed outliers; $M = 1104$ ms, S.D. = 418 ms) also showed no significant effect of ratio ($R = 0.21$, $p = 0.65$; intercept = 934, slope = 106).

2.6.2.3. Comparison to healthy participants. The patient's results were significantly worse than healthy participants' concerning overall mean accuracy, and accuracy with ratios 1.5 and 2 (see Table 3 for all results of this section, and also Fig. 3B). There was overall no difference in accuracy for the most difficult condition (ratio 1.3). Most importantly, there was no difference in the distance effect between the patient and the healthy participants (difference in accuracy between the smallest and largest ratio).[7] The patient's overall correct RTs did not significantly differ from healthy participants'.

2.6.3. Digit comparison

2.6.3.1. Method. We tested underlying numerical quantity representation through digit comparison which does not involve non-symbolic stimuli, therefore testing the quantity system through another entry. Typically, responses become faster as the distance between the digits to be compared increases (distance effect); this is thought to reflect decrease of overlap of underlying numerical representations, similarly to the distance effect on accuracy scores in the previous tasks. In this task, all possible combinations of digits 1–9 were used to create 36 pairs of digits. The distance between the digits constituting different pairs therefore varied (from 1 to 8). The patient was instructed to respond as accurately and as quickly as possible, by pressing the mouse button on the same side as the digit representing the larger quantity (using his left or right index). Before the test began, 10 training trials with feedback were administered.[8] Each test trial started with the presentation of the pair of white digits on a black background (duration = 700 ms), each digit on either side of a fixation circle. If the patient had not responded during the presentation of the digits, the fixation remained until he responded. The patient performed the test trials over 2 blocks of 36 trials (total of 72 trials). The pairs of digits were each presented twice (larger quantity presented once left and once right of fixation for each pair, randomized across trials and across blocks). Accuracy and RTs were measured. RTs were analyzed in relation to the distance between digits to detect a distance effect.

2.6.3.2. Patient's results[9]. Overall accuracy was good (98% correct). Mean correct RT (computed after having discarded outliers) was 681 ms (S.D. = 84) (see Fig. 3C). Correct RTs tended to decrease across distance, although this effect did not reach significance ($R = -1.62$, $p = 0.11$; intercept = 726; slope = −12; difference in RTs between the smallest and largest distances = 60 ms). Correct RTs also tended to correlate negatively with distance ($r = -0.22$, $p = 0.11$).

2.6.3.3. Comparison to healthy participants. The patient's overall accuracy did not significantly differ from the healthy participants' (see Table 3 for all results of this section, and also Fig. 3C). The patient's overall correct RT was significantly slower than healthy participants', and the intercept of regression of correct RTs against distance was higher than healthy participants', also indicating a slower performance. However, importantly, the patient's distance effect did not significantly differ from healthy participants', either when measured as the difference in RTs between the smallest and largest distance, the slope of the regression of correct RTs against distance, or by the correlation between correct RTs and distance.

2.6.4. Number-size Stroop digit comparison

2.6.4.1. Method. In this task we tested whether Arabic digits elicited an automatic access to numerical quantity in the patient. Pairs of digits (1–7, 1–8, 2–7, 2–9, 3–8

[6] The patient probably did not show a ratio effect because the ratios tested here were too easy. Data from the "Dots comparison" task (Section 2.6.1. and Fig. 3A) support this hypothesis, as the healthy participants (and the patient) are essentially at ceiling with ratio 1.3 in this task.

[7] A closer look at the healthy participants' single-subject data revealed that 73% of them showed a difference in accuracy between the smallest and largest ratio in the direction of a distance effect, but χ^2 individual analyses showed that this difference was significant for only one subject. The other 27% of healthy participants showed a null difference, as the patient did, or, in one case, an inverse distance effect. These additional results bring support to the hypothesis that the patient did not show a distance effect in this task because the ratios were too easy.

[8] Healthy participants only performed five training trials.

[9] Extremes (distances 1 and 8) were excluded before computing the distance effect on correct RTs, because of anchoring effects.

Table 3
Comparison of patient's vs. healthy participants' results in the four tasks tapping into the semantic representation of numbers

	Patient	Healthy participants		*t*-value (d.f. = 14)	*p*-value (two-tailed)
		Mean	S.D.		
Dots comparison (d.f. = 17)					
Accuracy (%)					
Overall	87	81	11	0.53	0.60
Regression of accuracy against ratio					
Slope	1.23	1.10	0.24	NA	Within 2S.D.s
Correlation with ratio	0.84	0.83	0.11	−0.11	0.91
w (estimate of the internal Weber Fraction)	0.14	0.17	0.05	−0.58	0.57
Dots addition and comparison					
Accuracy (%)					
Overall*	88	95	3	−2.26	<0.05
Constant area	96	95	3	0.32	0.75
Constant density**	79	96	5	−3.29	<0.01
Ratio 1.3	88	89	10	−0.10	0.92
Constant area	88	89	9	−0.11	0.92
Constant density	88	88	13	0	1.00
Ratio 1.5**	88	98	3	−3.23	<0.01
Constant area	100	98	5	0.39	0.70
Constant density**	75	98	4	−5.57	<0.01
Ratio 2**	88	100	2	−5.81	<0.01
Constant area	100	99	3	0.32	0.75
Constant density**	75	100	0	–	–
Difference ratio 2 – ratio 1.3	0	11	10	−1.07	0.31
Constant area	12	10	11	0.18	0.86
Constant density	−4	12	13	−1.19	0.25
RTs (ms)					
Overall	1104	896	222	0.91	0.38
Digit comparison					
Accuracy (%)					
Overall	98	99	1	−0.97	0.35
RT (ms)					
Overall**	681	484	43	4.46	<0.01
Distance 2**	684	507	56	3.06	<0.01
Distance 7**	624	467	30	5.07	<0.01
Difference distance 2–7	60	40	47	0.41	0.69
Regression of RTs against distance					
Intercept*	726	524	70	2.80	<0.05
Slope	−12	−10	8.24	NA	Within 2S.D.s
Correlation of RT with distance	−0.22	−0.23	0.19	0.08	0.94
Number-size Stroop digit comparison					
Accuracy (%)					
Overall	93	96	3	−0.97	0.35
Congruent	100	98	4	0.48	0.64
Incongruent	86	93	5	−1.36	0.20
Incongruent– congruent	−14	−5	7	1.25	0.23
RT (ms)					
Overall	802	574	201	1.10	0.29
Congruent	768	537	154	1.45	0.17
Incongruent	846	609	156	1.47	0.16
Incongruent–congruent	78	72	69	0.09	0.93

NA: statistical analysis was not possible due to differences among the healthy participants' error variances.
* Patient significantly differs from healthy participants' at $p < 0.05$.
** At $p < 0.01$.

and 3–9, distance of 5, 6 or 7) were presented and the patient had to judge the physical size of the digits, indicating as accurately and as fast as possible which digit was physically bigger, by pressing on the corresponding mouse button (using his left and right indexes). Numerical size of digits was to be ignored, and was congruent with physical size on half the trials, and incongruent on the other half (randomization across trials). Typically, RTs are slower in the incongruent condition compared to the congruent condition if numerical quantity is automatically accessed by the perception of the digit (number-size incongruence effect). In half the trials the physically larger digit was on the left. The numerically larger digit was also on the left on half the trials. Before the test began, six training trials with feedback were administered. Each trial started with the presentation of a pair of digits (700 ms; digits separated by a distance varying from 4.5° to 5°), each digit on either side of a fixation circle (white on a black background, visual angle of 2°). The fixation remained for 300 ms after the digits disappeared, and longer (1500 ms) if no response had been detected. The digits differed by either 8, 16, 22, 30 or 38 units of character size, their visual angle varying from 0.8° to 2.1° of height and from 0.4° to 1.3° of width. The patient performed 56 test trials in one block. Accuracy and RTs were measured, and RTs were analyzed in relation to the condition (congruent/incongruent) to detect an incongruence effect.

Table 4
Comparison of patient's vs. healthy participants' results in verbal numerosity estimation with external calibration

Response	Patient	Healthy participants		t-value (d.f. = 14)	p-value (two-tailed)
		Mean	S.D.		
Mean	61.67	46.91	7.02	2.04	0.06
Constant area*	63.96	45.20	8.08	2.25	<0.05
Constant density	59.37	48.62	7.00	1.49	0.16
Numerosity-response correlation coefficient	0.74	0.88	0.05	−1.64	0.13
Constant area	0.69	0.86	0.07	−1.47	0.16
Constant density*	0.79	0.92	0.03	−2.77	<0.05
Mean variation coefficient**	0.43	0.23	0.05	3.87	<0.01
Constant area**	0.45	0.20	0.05	4.84	<0.01
Constant density*	0.34	0.20	0.06	2.26	<0.05
RT (ms)					
Overall	4396	2773	794	1.98	0.07

* Patient significantly differs from healthy participants' at $p < 0.05$.
** At $p < 0.01$.

2.6.4.2. Patient's results. Overall accuracy was good (93% correct), and was lower in the incongruent condition (86%) than in the congruent condition (100%), although this difference did not reach significance (difference = −14%; $\chi^2(1) = 2.42$, $p = 0.12$). Mean correct RT (computed after having discarded outliers) was 802 ms (S.D. = 147 ms). Correct RTs were slower in the incongruent condition (846 ms, vs. 768 ms in the congruent condition; difference incongruent − congruent = 78 ms); this effect approached statistical significance ($t(48) = 1.93$, $p = 0.06$) (see Fig. 3D). The patient therefore presented an incongruence effect, although it did not reach significance.

2.6.4.3. Comparison to healthy participants. The patient did not differ from healthy participants on overall accuracy, accuracy for the congruent and incongruent conditions separately, nor on the difference in accuracy between the incongruent and congruent conditions (see Table 3 for all results from this section, and also Fig. 3D). This was also the case for the same comparisons on RTs. There was therefore no significant difference in incongruence effect. These results suggest intact automatic access to numerical quantity.

In a mirror task in which the patient was to judge digits on their numerical size, and ignore physical size, the patient's effect of interference from physical size was also comparable to healthy participants' on both accuracy and RT scores (although the patient's overall RTs, and RTs for each condition were slower than healthy participants').

2.6.5. Comment on tasks tapping into the semantic representation of numbers

In sum, although the patient differed from the healthy participants on some quantitative measures (overall RTs in digit comparison; some accuracy measures in dots addition and comparison), his performance as regards qualitative measures was always comparable to healthy participants (distance effects in dots comparison, dots addition and comparison, and digit comparison; number-size incongruence effect in the Stroop digit comparison). Moreover, the patient's performance was comparable to healthy participants' on all measures of two out of the four tests, clearly suggesting intact underlying numerical representation (dots comparison), and automatic access to numerical representation from Arabic digits (number-size Stroop digit comparison).

We next tested the level of external calibration by administering the first estimation task with external calibration, which means showing examples of correct responses, to see if the patient would be able to take these into account to adjust his responses. We hypothesized that the patient would not benefit from external calibration as much as healthy participants, as it is thought to implicate a strategic component, which might be impaired in relation to the patient's executive deficits.

2.7. Verbal numerosity estimation with external calibration

2.7.1. Method

The stimuli and test procedure were exactly the same as in first experiment (see Section 2.5.1), except that each block was preceded by external calibration, which consisted of examples of stimuli other than those tested, but sampling the same range (numerosities 15, 60 and 140). Two examples of each calibration numerosity were presented: one from a set of constant total occupied area, and the other from a set of constant density. The patient was informed of the exact numerosity (e.g.: "Here are 15 dots"). Calibration dots remained on the screen for 10 s or less, if the patient was ready sooner to see the next set. In the next description of results sections, we compare the performance on this task with that obtained in the first experiment.

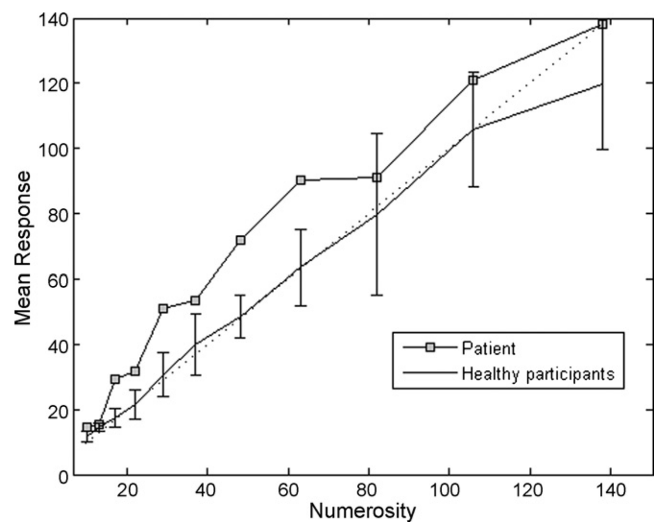

Fig. 4. Patient's vs. healthy participants' performance in verbal numerosity estimation with external calibration. (Error bars represent ±1 standard deviation; the dashed line represents theoretical exact correspondence between numerosity and mean response; mean responses can therefore be visualized as under- or overestimates if they fall respectively below or above the dashed line; note difference in mean response scale when comparing to Fig. 2.)

2.7.2. Patient's results[10]

Compared to results from the first experiment, and concerning extreme answers, the patient's responses ($M = 61.67$, S.D. = 34.66; see Fig. 4) remained consistently larger than the correct response across numerosities, although this overestimation was much less pronounced (responses ranged from 13 to 160, at the most 170 for numerosity 138; compare Fig. 4 to Fig. 2). Performance was identical to the first experiment as concerns the presence of signatures of estimation processes. There was again a numerosity-response correlation ($r = 0.74$, $p < 0.01$; same value) and scalar variability (again a stable variation coefficient across numerosities: $R = 1.71$, $p = 0.21$; intercept = 0.32, slope = 0.002). Moreover, RTs (computed after having removed outliers; $M = 4396$ ms, S.D. = 564 ms), which were more than one second faster than in the first experiment, were again flat across numerosities ($R = 2.34$, $p = 0.13$; intercept = 3989, slope = 9). Finally, the mean variation coefficient remained high ($M = 0.43$, S.D. = 0.15; same value).

2.7.3. Comparison to healthy participants

As concerns extreme answers, the patient's mean response was still higher than healthy participants'. However, this difference was 3–4 times smaller than in the first experiment and no longer reached statistical significance (see Table 4 for all results of this section; compare to Table 2). Also, mean response calculated for each numeros-

[10] Unless specified otherwise, we report results and analyses excluding data from the extremes numerosities (10 and 138) to avoid noise from anchoring effects.

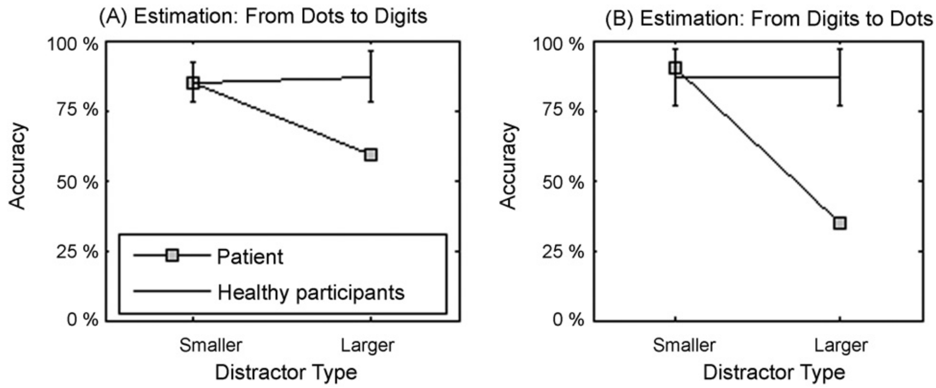

Fig. 5. Comparison of the patient's vs. healthy participants' performance in the two forced-choice estimation tasks. (A) Distractor effect on accuracy in the "Dots to digits" task; (B) distractor effect on accuracy in the "Digits to dots" task.

ity was over 2 standard deviations of healthy participants' for some numerosities only (5 out of 9; see Fig. 4) as opposed to all of them in the first experiment (compare to Fig. 2). Therefore, overestimation was overall reduced in this task with external calibration, in comparison to the first experiment. Regarding signatures of estimation processes, performance was also improved in this task as concerns the quantitative value of the numerosity-response correlation, as the patient no longer statistically differed from healthy participants' on this measure (although his correlation still remained lower than healthy participants'). However, the patient's mean variation coefficient remained significantly higher than healthy participants'. The patient's mean RT remained slower than healthy participants' although this difference no longer reached statistical significance (previously significant).

In sum, in comparison to the first experiment, the patient's performance again showed typical signatures of an estimation process: his responses correlated with numerosity and respected scalar variability, and his RTs were flat across numerosities. In addition, the patient's performance improved on some quantitative measures, as he presented less overestimation and was no longer statistically slower than healthy participants. Results therefore suggest that the patient had difficulties calibrating himself in the first task, and benefited somewhat from the external calibration to counter overestimation in this task. However, estimation precision, as measured by the variation coefficient, was still lower than healthy participants', probably due to a greater variation in response, and also due to the mean response still being somewhat higher than the healthy participants'.

In order to rule out the possibility of a deficit situated at the level of verbal output, we designed two forced-choice estimation tasks. The first one was similar to the main estimation task, in that the stimuli were the same clouds of dots; the output changed from number words to Arabic digits, and the forced-choice design allowed us to investigate whether the deficit was not limited to excessive output variability. In the second task, both input and output changed (presenting Arabic digits as stimuli and clouds of dots as responses to choose from) and again allowed testing of another type of output, but this time non-symbolic. We hypothesized that performance would remain disrupted in these two new tasks, as we supposed the deficit to be independent from type of output.

2.8. Forced-choice estimation "from dots to digits"

2.8.1. Method

We presented the same stimuli as in the first estimation task (see Section 2.5.1), and asked the patient to choose, as accurately and as fast as possible, the corresponding Arabic digit among two choices (the correct response and a distractor). The distractor was smaller (ranging from 4 to 62) on half the trials, and larger (ranging from 22 to 309) on the other half. The ratio between the correct response and the distractor was maintained constant (~2.24), and was picked to match the degree of overestimation in the first task. The patient indicated the correct digit by pressing on the corresponding mouse button (left or right, using his left or right index). Dots were presented as in the first experiment, but were followed by an empty white central disc under which two white digits appeared: one on the left and the other on the right, until the patient responded. The digits then disappeared and 700 ms elapsed before the next trial. The patient performed a total of 132 trials over three blocks. On half the trials, the correct response was presented on the left, and vice-versa for the other half. Each numerosity (cloud of dots), was presented 12 times (half the time with a smaller distractor). We compared accuracy scores from each condition (smaller/larger distractor) to detect consistent over- or underestimation.

2.8.2. Patient's results

The patient's overall accuracy was of 72% correct, and was significantly lower in the condition with a larger distractor (59%) compared to the condition with a smaller distractor (85%) (difference = 26%; $\chi^2(1) = 9.61$, $p < 0.01$) (see Fig. 5A). His performance was significantly above chance in the condition with a smaller distractor ($\chi^2(1) = 32.06$, $p < 0.01$), but not in the condition with a larger distractor ($\chi^2(1) = 2.18$, $p = 0.14$). Correct RTs (computed after having removed outliers; $M = 1934$ ms, S.D. = 606 ms) were not significantly influenced by distractor type ($t(69) = -0.60$, $p = 0.55$).

2.8.3. Comparison to healthy participants

The patient significantly differed from healthy participants on overall accuracy. His accuracy was significantly lower than the healthy participants' but only in the condition with larger distractors and not with smaller ones (see Table 5 for all results of this section, and also Fig. 5A). The difference in accuracy between the two conditions was higher in the patient compared to healthy participants, but this did not reach significance. Finally, the patient's mean overall RT was significantly slower than the healthy participants'. Healthy participants' RTs were also analyzed and similarly to the patient, there was no effect of distractor type ($t(58) = -0.86$, $p = 0.40$).

In sum, the results point to a deficit which is not limited to number words or to excessive response variability, but which also persists in a forced-choice paradigm, and only with larger distractors, in line with the overestimation found in the main estimation task. This suggests that the deficit is not situated at the output level.

2.9. Forced-choice estimation "from digits to dots"

2.9.1. Method

In order to find out whether the patient's estimation deficit would also occur with non-symbolic output, we administered a forced-choice estimation task which mirrors the previous one, presenting an Arabic digit and asking the patient to choose as accurately and as quickly as possible the corresponding cloud of dots among two choices (the correct response and a distractor). Procedure was the same as in the previous task. However, we used fewer stimuli as in the previous task, presenting only four digits (29, 48, 82 and 138) and their corresponding distractors (chosen as described for the previous task). The dots' diameter subtended a visual angle varying from 0.1° to 0.3° and the white discs' diameter a visual angle of 14° (distance between the discs = 0.4°). Each digit was presented 24 times (half the time with a smaller distractor). On half the trials the correct response appeared on the left. Half the sets of clouds were of constant dot size, and for the other half, the envelope of the area covered by the dots was held constant. Again, we compared accuracy scores from each condition (smaller/larger distractor) to detect consistent over- or underestimation.

2.9.2. Patient's results

The patient's overall accuracy was of 63% correct, and was significantly lower in the condition with a larger distractor (35%) compared to the condition with a smaller distractor (90%) (difference = 55%; $\chi^2(1) = 27.78$, $p < 0.01$) (see Fig. 5B). His performance was significantly above chance in the condition with a smaller distractor ($\chi^2(1) = 30.08$, $p < 0.01$), but was significantly worse than chance in the condition with a larger distractor ($\chi^2(1) = 4.08$, $p < 0.05$), indicating a clear bias to select the larger set of dots. Correct RTs (computed after having removed outliers; $M = 1877$ ms, S.D. = 641 ms) were not significantly influenced by distractor type ($t(55) = 0.60$, $p = 0.55$).

2.9.3. Comparison to healthy participants

The patient significantly differed from healthy participants on overall accuracy, accuracy being significantly lower than healthy participants' only in the condition with larger distractors and not with smaller ones (see Table 5 for all results of this section, and also Fig. 5B). Also, the difference in accuracy between the two conditions

Table 5
Comparison of patient's vs. healthy participants' results in the forced-choice estimation tasks

	Patient	Healthy participants		*t*-value (d.f. = 14)	*p*-value (two-tailed)
		Mean	S.D.		
From dots to digits					
Accuracy (%)					
Overall**	72	86	4	−3.39	<0.01
Smaller distractor	85	85	7	–	–
Larger distractor **	59	87	9	−3.01	<0.01
Smaller – larger distractor	26	−0.02	0.14	1.94	0.07
RT (ms)					
Overall*	1934	1122	288	2.73	<0.05
From digits to dots					
Accuracy (%)					
Overall**	63	87	6	−3.87	<0.01
Smaller distractor	90	87	10	0.29	0.78
Larger distractor**	35	87	10	−5.04	<0.01
Smaller – Larger distractor**	55	1	17	3.08	<0.01
RT (ms)					
Overall	1877	1307	466	1.18	0.26

* Patient significantly differs from healthy participants' at $p < 0.05$.
** At $p < 0.01$.

was significantly higher in the patient compared to healthy participants. Finally, the patient's mean overall RT was not significantly different from healthy participants'. Healthy participants' RTs were also analyzed and similarly to the patient, there was no effect of distractor type ($t(58) = -0.20$, $p = 0.84$).

In sum, these results show that the deficit was not limited to symbolic output, but extended to non-symbolic output, and indicate a bias to select the larger response. Indeed, the patient consistently picked the larger set of dots, therefore matching large sets to smaller Arabic digits. If given quantities of dots were consistently linked to larger digits, as suggested by the overestimation in tasks with dots as stimuli and number words (see Section 2.5) or digits (see Section 2.8) as output, the patient would have systematically picked the smaller set of dots in this last task. This argues against a deficit at the level of the semantic representation of numbers, and for impairment at the level of translation from representation to output, which generalizes to different types of output.

2.10. Effect of number range?

The patient's overestimation deficit in the estimation task without external calibration was particularly striking with larger numerosities (63–106). We wished to further explore the possibility that the patient's estimation deficit was limited to or more pronounced with larger values. We therefore conducted additional testing, comparing the patient's performance to that of healthy participants separately over small (<50) and large (>50) number ranges. This was carried out in all estimation tasks (estimation without external calibration, estimation with external calibration, forced-choice estimation "from dots to digits" and forced-choice estimation "from digits to dots") and the results are reported in Table 6.

In the first two estimation tasks, we conducted linear regressions of response against numerosity separately over each range, obtaining slopes which can be interpreted as indicating underestimation (<1) or overestimation (>1), for the patient and for each healthy participant. In the first estimation task (without external calibration), the patient's slope was larger than healthy participants' over both ranges[11]; however, this was more marked over the large range. In the second estimation task (with external calibration), the patient's slope was comparable to healthy participants' over both ranges, indicating spared performance.[12] For these two tasks, we also conducted comparisons of absolute differences in slope, within each task (difference in slope due to range), and across tasks (difference in slope due to calibration, within each number range). Results show that in both estimation tasks, the patient's difference in slope between the small and large ranges was significantly larger than healthy participants', suggesting that number range influenced his performance more than it influenced that of the healthy participants. Regarding differences in slope due to external calibration, looking separately at small and large ranges, the patient did not differ from the healthy participants over the small range, but presented a dramatically (and significantly) higher difference in slope than the healthy participants over the large range. This suggests that the patient's overestimation was more pronounced in the large range to start with, and that it was more susceptible to calibration in this range. In the forced-choice estimation tasks, we calculated accuracy separately over each range for each distractor type (smaller or larger). In the forced-choice estimation task "from dots to digits", the patient's accuracy score was comparable to healthy participants' over both ranges when the distractor was smaller, consistent with previous analyses showing spared performance in this condition. However, when the distractor was larger, the patient's performance was significantly worse than healthy participants' only over the large range. A similar pattern emerged in the other forced-choice estimation task ("from digits to dots"), showing again spared performance over both ranges with smaller distractors, and a more pronounced deficit over the large range compared to the small range with larger distractors. In sum, results from the different estimation tasks converged to suggest that the patient's estimation deficit was more severe over the larger number range (>50), although not limited to it. Also, results showed that calibration had a clearly greater impact in the larger number range, bringing more evidence that the source of the patient's overestimation bias seems to be higher-order.

2.11. Sensitivity to non-numerical parameters

When compared to the healthy participants, the patient tended to show a greater sensitivity to non-numerical parameters, which probably reflects impairment in resolving conflicts between dimensions. In this section we report analyses focused on detecting a possible influence of non-numerical parameters in five of the tasks that were administered to the patient (all four estimation tasks, as well as the dots addition and comparison task). For some tasks, ANOVAs were performed, which incorporated a factor accounting for the different types of controls performed on stimuli.

In the "estimation without external calibration" task, the patient overestimated the stimuli in comparison to the healthy participants, similarly for both kinds of displays (where either area or density was kept constant across numerosities; see Table 2). However, a 9×2 ANOVA with numerosity (13–106) and non-numerical parameter (constant area or constant density of dots) as variables showed that the patient gave larger responses for the trials of constant density ($M = 114.26$, S.D. = 113.08) in comparison to the trials of constant area ($M = 62.11$, S.D. = 38.39) ($F(1, 36) = 13.23$, $p < 0.01$), especially in the large number range (interaction between numerosity and non-numerical parameter: $F(8, 36) = 3.32$, $p < 0.01$). In contrast with the patient's results, there was no main effect of non-numerical parameter ($F(1, 252) = 0.99$, $p = 0.32$) or interaction with numerosity ($F(8, 252) = 0.45$, $p = 0.89$) in the responses of healthy participants.

This tendency reversed in the "estimation with external calibration" task, in which healthy participants showed a similar (but much less marked) effect, giving larger responses in the trials of constant density ($F(1, 252) = 3.78$, $p = 0.05$), particularly in the large number range ($F(8, 252) = 3.82$, $p < 0.01$). In contrast, no such effect was present in the patient's responses in this task (main effect of non-numerical parameter: $F(1, 36) = 0.27$, $p = 0.61$; interaction with numerosity: $F(8, 36) = 0.44$, $p = 0.89$). The comparison of the two tasks suggests that the patient's sensitivity to non-numerical parameters observed for the estimation task without external calibration represents an exaggeration of a tendency which is present in the whole population. This tendency, while above average in some tasks (esti-

[11] Following Crawford and Garthwaite (2004), we wished to statistically compare the slope of the patient's regression to that of healthy participants'. Given that there were differences among the healthy participants' error variances, this test was not applicable and we instead determined whether the patient's slope was within 2 S.D.s of healthy participants' slopes.

[12] Id as in previous footnote.

Table 6
Separate comparison by number range (small or large) of patient's vs. healthy participants' results in the four different estimation tasks

	Patient	Healthy participants		t-value (d.f. = 14)	p-value (two-tailed)
		Mean	S.D.		
Estimation without external calibration (response)					
Slope over small number range (13–48)[a]	1.65	0.77	0.32	NA	Over 2S.D.s
Slope over large number range (63–106)[b]	2.77	0.56	0.51	NA	Over 4S.D.s
Absolute difference in slope (small – large)**	1.12	0.34	0.24	3.15	<0.01
Estimation with external calibration (response)					
Slope over small number range (13–48)	1.52	1.01	0.27	NA	Within 2S.D.s
Slope over large number range (63–106)	0.74	0.99	0.33	NA	Within 2S.D.s
Absolute difference in slope (small – large)*	0.78	0.28	0.20	2.42	<0.05
Effect of external calibration within each number range (response)					
In the small number range (13–48)					
Absolute difference in slope (without calibration – with calibration)	0.13	0.31	0.25	−0.70	0.50
In the large number range (63–106)					
Absolute difference in slope (without calibration – with calibration)**	2.03	0.66	0.38	3.49	<0.01
From dots to digits (accuracy, %)					
Smaller distractor					
Over small number range (13–48)	94	93	5	0.19	0.85
Over large number range (63–106)	61	71	18	−0.54	0.60
Larger distractor					
Over small number range (13–48)	50	82	15	−2.07	0.06
Over large number range (63–106)**	67	95	5	−5.42	<0.01
From digits to dots (accuracy, %)					
Smaller distractor					
Over small number range (29 and 48)	88	96	6	−1.29	0.22
Over large number range (82 and 138)	92	78	18	0.75	0.46
Larger distractor					
Over small number range (29 and 48)*	58	88	10	−2.91	<0.05
Over large number range (82 and 138)**	13	85	15	−4.65	<0.01

NA: statistical analysis was not possible due to differences among the healthy participants' error variances.
[a] Patient's slope exceeds healthy participants' by 2 S.D.s.
[b] By 4S.D.s.
* Patient significantly differs from healthy participants' at $p < 0.05$.
** At $p < 0.01$.

mation without external calibration), is reduced in other contexts, such as in the estimation task with external calibration. It also disappeared in both forced-choice estimation tasks, in which no effect of non-numerical parameter was found in the patient's accuracy ("from dots to digits": $\chi^2(1) = 1.84$, $p = 0.18$; "from digits to dots": $\chi^2(1) = 0.18$, $p = 0.67$) or in the healthy participants' accuracy for the "from digits to dots" task ($t(28) = -0.60$, $p = 0.56$). However, healthy participants were influenced by non-numerical parameters in the "from dots to digits" task ($t(28) = -2.56$, $p < 0.05$), showing a slightly higher accuracy in trials of constant density ($M = 89\%$, S.D. = 5%; constant area: $M = 84\%$, S.D. = 5%). Furthermore, as a result of the discrepancy between experiments in the patient's and the healthy participants' sensitivity to non-numerical parameters, some differences arose in the direct comparison of the patient's and the healthy participants' responses (see Table 4). However, it is not clear whether these differences emanate from the patient changing his sensitivity to non-numerical parameters between the two tasks, or from the healthy participants doing so.

Finally, in the dots addition and comparison task the patient showed again a tendency to be affected by non-numerical parameters. Indeed, his performance was comparable to healthy participants' in trials of constant area, whereas it was significantly worse in trials of constant density for some measures (overall accuracy, accuracy with ratio 1.5 and accuracy with ratio 2; see Table 3).

In sum, the patient was influenced by non-numerical parameters in all tasks except the two last forced-choice tasks. His sensitivity to non-numerical parameters was inconsistent across tasks or even sometimes within task, as he showed a deficit with trials of constant density for some measures and impairment with trials of constant area for others. It is however interesting to note that the patient's tendency to give more overestimated responses to our least dense stimuli in the first estimation task (without external calibration) is in accordance with a known effect of density on estimation in healthy subjects: sets of dots lead to smaller estimates when their density is higher, compared to when they are more spread out (Hollingsworth, Simmons, Coates, & Cross, 1991; Krueger, 1972). In fact, the healthy participants showed this effect in the estimation task with external calibration. The patient's larger estimates in trials of constant density (increasing area as numerosity increases) compared to trials of constant area (increasing density as numerosity increases) could therefore perhaps be explained as an exaggeration of a normal tendency. However, this cannot explain the patient's extreme overestimation in both types of trials. In conclusion, results from the different tasks suggest difficulties in focusing on numerosity and in ignoring continuous non-numerical parameters.

3. Discussion

This study reports the case of a patient with focal frontal lobe damage who presents a cognitive estimation deficit in various quantitative domains (size, weight, numerosity and time). In addition, we found the patient to present a neuropsychological profile compatible with lesion localization, general posterior functions and general intellectual abilities being spared in the context of isolated executive and attention impairments. This type of report is similar to previous ones associating a cognitive estimation deficit to impaired selective and regulative processes sub-served by the frontal lobes (Della Sala et al., 2004; Shallice & Evans, 1978; Smith & Milner, 1984, Experiment 1).

In this study, we further determined that verbal numerosity estimation (which draws more strongly on perceptual processes than cognitive estimation) was also clearly impaired in this patient. This impairment was characterized by extreme answers (overestimation), but also by a larger variability in response, and a tendency to be influenced by non-numerical continuous parameters that covaried with numerosity (such as the size of the area occupied by the set of stimuli, or the density of the set of stimuli). However, this patient's verbal numerosity estimation process was not completely impaired, as it showed typical signatures of estimation processes, in particular, scalar variability (Dehaene & Marques, 2002;

Gallistel & Gelman, 1992; Izard & Dehaene, 2008; Whalen et al., 1999).

Although this case might seem in contrast with previous reports suggesting a link between parietal damage and verbal numerosity estimation deficits (Delazer et al., 2006; Warrington & James, 1967), we suggest it is complementary, in that the source of the verbal numerosity deficit probably differs (executive processes vs. semantic representation of numbers), as we will develop below.

In accord with our prediction, the patient's verbal numerosity estimation deficit did not seem to reflect impairment at the level of the semantic representation of numbers. Indeed, several tests suggested general sparing of numerical abilities, consistent with the sparing at the anatomical level of the parietal lobes which are known to play an important role in numerical representation (for a review, see Dehaene et al., 2003). Although the patient's performance was intact in these tasks as regards qualitative measures, he differed from healthy participants on some quantitative measures, being significantly slower in one task, and less accurate in another. We believe that these differences may be linked to attention fluctuations, which were present in the neuropsychological examination. It has indeed been shown that slowing and excessive variation of RTs can reflect frequent lapses of attention and an instability of attention performance (Benke, Delazer, Bartha, & Auer, 2003). Also, it might come as a surprise that the patient's performance was intact in one of the tasks tapping into the semantic representation of numbers (the number-size Stroop task), given the inhibition difficulties that he presented in the main neuropsychological investigation. We speculate that this could again perhaps be linked to attention fluctuations, which might be able to account for fluctuations in inhibition skills across tasks.

We had also predicted that external calibration could be impaired, as a recent study has suggested possible involvement of strategic processes in external calibration (Izard & Dehaene, 2008). When externally calibrated, the patient was able to adjust his responses to some extent (less overestimation), suggesting some sparing of the capacity to draw inferences from external reference. He still presented some overestimation and a large variability in responses, as well as still being somewhat influenced by non-numerical parameters. We do not rule out the hypothesis that another patient with more pronounced executive difficulties may not benefit at all from external calibration.

In accord with our last hypothesis, results suggested that the deficit was not situated at the level of verbal output. Indeed, we established that the estimation deficit was not limited to excessive variability in responses, as it persisted in a forced-choice paradigm. Also, we determined that it was not specific to the verbal output modality, as the patient also presented marked difficulties with another type of symbolic output and with non-symbolic output. Interestingly, this last paradigm also brought evidence that the patient's estimation deficit was not a consistent erroneous link between the semantic representation of numbers and responses, which might have suggested impairment at the semantic representation level. Indeed, although the patient linked clouds of dots to larger digits in the other estimation tasks, in this last estimation paradigm, he consistently linked digits to the larger set of dots, rather than the smaller one, thus breaking the overestimation pattern. It therefore seems that the patient presented a bias toward selecting large quantities, whether number words, digits, or sets of dots. Taken together with the fact that the level of the semantic representation of numbers was intact, these results therefore clearly point to impairment at the level of translation from representation to output, as we had hypothesized.

Interestingly, further analyses carried out over several tasks revealed that the estimation deficit was more marked with (but not limited to) stimuli drawn from the large number range (>50) as opposed to the small number range (<50). This could be a difficulty effect, as we are much more familiar with smaller quantities, and it may therefore be easier to judge their approximate quantity than larger ones. Some of the tasks tapping into the semantic level only presented stimuli from a very small number range (1–9). We therefore cannot rule out the possibility that performance on these tasks was spared because they involved only small, over-learned numbers. However, larger stimuli were used in the other tasks tapping into the semantic representation, for which the patient's performance was comparable to healthy participants', suggesting sparing of the semantic representation of these larger numbers.

Finally, as concerns the influence of non-numerical parameters, it is interesting to note that the patient's performance presented similarities to that of healthy participants, indicating that his difficulties focusing on numerosity might be an exaggeration of a normal effect, and therefore bringing an additional argument against a fundamental numerical deficit. Indeed, it is known that sets of dots lead to smaller estimates when their density is higher, compared to when they are more spread out (Hollingsworth et al., 1991; Krueger, 1972), and the healthy participants were found to be influenced in this way in one of the estimation tasks, over larger numerosities only.

We conclude that impairment of verbal numerosity estimation can occur in the absence of clear deficits in the semantic representation of numbers, similarly to reports of cognitive estimation deficits in the absence of impairment of general semantic knowledge (Della Sala et al., 2004; Shallice & Evans, 1978; Smith & Milner, 1984, Experiment 1). We speculate that the cognitive estimation and verbal numerosity estimation deficits presented by this patient both stem from executive deficits disrupting the translation from semantic representation (of general knowledge or of numbers) to output, in relation to his frontal lesions.

Acknowledgments

We warmly thank the patient and the healthy participants for participating in this study, as well as Sandra Barcelona-Lehmann, Linda Revkin and two anonymous reviewers for useful comments and suggestions on previous versions of this manuscript. This study was supported by INSERM, CEA, two Marie Curie fellowships of the European Community (MRTN-CT-2003-504927; QLK6-CT-2002-51635), and an Austrian FWF grant (P18896-B05).

References

Aschenbrenner, S., Tucha, O., & Lange, K. W. (2001). *Regensburger Wortflüssigkeits-Test*. Göttingen Bern Toronto Seattle: Hogrefe Verlag für Psychologie.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*(Pt 11), 2189–2202.

Benke, T., Delazer, M., Bartha, L., & Auer, A. (2003). Basal ganglia lesions and the theory of fronto-subcortical loops: Neuropsychological findings in two patients with left caudate lesions. *Neurocase*, *9*, 70–85.

Brand, M., Kalbe, E., & Kessler, J. (2002a). Qualitative and quantitative differences in cognitive estimation of patients with probable Alzheimer's disease from healthy controls – what are the differences? In P. Calabrese & A. Neugebauer (Eds.), *Memory and emotion* (pp. 425–429). New Jersey: World Scientific.

Brand, M., Kalbe, E., & Kessler, J. (2002b). *Test zum kognitiven Schätzen (TKS)*. Göttingen: Hogrefe.

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *Public Library of Science Biology*, *4*, e125.

Crawford, J. R., & Garthwaite, P. H. (2004). Statistical methods for single-case studies in neuropsychology: Comparing the slope of a patient s regression line with those of a control sample. *Cortex*, *40*, 533–548.

Crawford, J. R., Garthwaite, P. H., Howell, D. C., & Venneri, A. (2003). Intra-individual measures of association in neuropsychology: Inferential methods for comparing a single case with a control or normative sample. *Journal of the International Neuropsychological Society*, *9*, 989–1000.

Dehaene, S., & Marques, J. F. (2002). Cognitive euroscience: Scalar variability in price estimation and the cognitive consequences of switching to the euro. *The Quarterly Journal of Experimental Psychology*, *55*, 705–731.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.

Delazer, M., Karner, E., Zamarian, L., Donnemiller, E., & Benke, T. (2006). Number processing in posterior cortical atrophy – a neuropsychological case study. *Neuropsychologia*, *44*, 36–51.

Della Sala, S., MacPherson, S. E., Phillips, L. H., Sacco, L., & Spinnler, H. (2004). The role of semantic knowledge on the cognitive estimation task – evidence from Alzheimer s disease and healthy adult aging. *Journal of Neurology*, *251*, 156–164.

Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB: A frontal assessment battery at bedside. *Neurology*, *55*, 1621–1626.

Flowers, K. A., & Robertson, C. (1985). The effect of Parkinson's disease on the ability to maintain a mental set. *Journal of Neurology, Neurosurgery, and Psychiatry*, *48*, 517–529.

Fox, A. M., Michie, P. T., Wynne, C. D., & Maybery, M. T. (2000). ERP correlates of response inhibition to elemental and configural stimuli in a negative patterning task. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *111*, 1045–1053.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74.

Helmstaedter, C., & Durwen, H. F. (1990). VLMT: Verbaler Lern- und Merkfähigkeitstest: Ein praktikables und differenziertes Instrumentarium zur Prüfung der verbalen Gedächtnisleistungen. *Schweizer Archiv für Neurologie und Psychiatrie*, *141*, 21–30.

Hollingsworth, W. H., Simmons, J. P., Coates, T., & Cross, H. A. (1991). Perceived numerosity as a function of array number, speed of array development, and density of array items. *Bulletin of the Psychonomic Society*, *29*, 448–450.

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*, 1221–1247.

Jackson, M., & Warrington, E. K. (1986). Arithmetic skills in patients with unilateral cerebral lesions. *Cortex*, *22*, 611–620.

Kohler, J., & Beck, U. (2004). Der Planungstest – Softwareprogramm zur Diagnostik von Planungsstörungen. EnJoiME Software (Beck and Kohler GbR). 78467 Konstanz.

Kohler, J., Beck, U., & Hohnecker, B. (2003). Neuropsychological assessment of basic planning competence. *Journal of the International Neuropsychological Society*, *9*, 519.

Krueger, L. E. (1972). Perceived numerosity. *Perception and Psychophysics*, *11*, 5–9.

Lehrl, S., Merz, J., Burkard, G., & Fischer, B. (1991). *Mehrfachwahl-Wortschatz Intelligenz Test*. Erlangen: Peri-Med Fachbuch.

Mendez, M. F., Doss, R. C., & Cherrier, M. M. (1998). Use of the cognitive estimations test to discriminate frontotemporal dementia from Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, *11*, 2–6.

Minturn, A. L., & Reese, T. W. (1951). The effect of differential reinforcement on the discrimination of visual number. *Journal of Psychology*, *31*, 201–231.

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*, 177–190.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*, 1708–1711.

Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 7457–7462.

Pesenti, M., Thioux, M., Samson, D., Bruyer, R., & Seron, X. (2000). Number processing and calculation in a case of visual agnosia. *Cortex*, *36*, 377–400.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*, 547–555.

Piazza, M., Mechelli, A., Price, C. J., & Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: An fMRI study. *Brain Research*, *1106*, 177–188.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*, 293–305.

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique. *Archives de Psychologie*, *28*, 286–340.

Roth, M., Schumacher, J., & Arnett, J. (2003). Die deutsche Version des Arnett Inventory of Sensation Seeking (AISS-D). In M. Roth & P. Hammelstein (Eds.), *Sensation Seeking: Konzeption, Diagnostik und Anwendung* (pp. 100–121). Göttingen: Hogrefe.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime Users Guide*. Pittsburgh: Psychology Software Tools Inc.

Shallice, T., & Evans, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex*, *14*, 294–303.

Smith, M. L., & Milner, B. (1984). Differential effects of frontal-lobe lesions on cognitive estimation and spatial memory. *Neuropsychologia*, *22*, 697–705.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.

Taylor, R., & O'Carroll, R. (1995). Cognitive estimation in neurological disorders. *British Journal of Clinical Psychology*, *34*(Pt 2), 223–228.

Warrington, E. K., & James, M. (1967). Tachistoscopic number estimation in patients with unilateral cerebral lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *30*, 468–474.

Wechsler, D. (1987). *WMS-R Wechlser Memory Scale-Revised*. San Antonio: Psychological Corporation.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Non-verbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*, 130–137.

Zimmermann, P., & Fimm, B. (2002). *Testbatterie zur Aufmerksamkeitsprüfung (TAP)*. Herzogenrath, Germany: Psychologische Testsysteme.