Non-Spatial Impairments Affect False-Positive Neglect Diagnosis Based on Cancellation Tasks

Hanne Huygelier[1] , Margaret Jane Moore[2], Nele Demeyere[2] and Céline R. Gillebert[1,2,*]

[1]Department of Brain and Cognition, KU Leuven, Leuven, Belgium
[2]Department of Experimental Psychology, University of Oxford, Oxford, UK

Abstract

Objective: To diagnose egocentric neglect after stroke, the spatial bias of performance on cancellation tasks is typically compared to a single cutoff. This standard procedure relies on the assumption that the measurement error of cancellation performance does not depend on non-spatial impairments affecting the total number of cancelled targets. Here we assessed the impact of this assumption on false-positive diagnoses. **Method:** We estimated false positives by simulating cancellation data using a binomial model. Performance was summarised by the difference in left and right cancelled targets (R-L) and the Centre of Cancellation (CoC). Diagnosis was based on a fixed cutoff versus cutoffs adjusted for the total number of cancelled targets and on single test performance versus unanimous or proportional agreement across multiple tests. Finally, we compared the simulation findings to empirical cancellation data acquired from 651 stroke patients. **Results:** Using a fixed cutoff, the rate of false positives depended on the total number of cancelled targets and ranged from 10% to 30% for R-L scores and from 10% to 90% for CoC scores. The rate of false positives increased even further when diagnosis was based on proportional agreement across multiple tests. Adjusted cutoffs and unanimous agreement across multiple tests were effective at controlling false positives. For empirical data, fixed versus adjusted cutoffs differ in estimation of neglect prevalence by 13%, and this difference was largest for patients with non-spatial impairments. **Conclusions:** Our findings demonstrate the importance of considering non-spatial impairments when diagnosing neglect based on cancellation performance.

Keywords: Perceptual disorders, task performance and analysis, psychometrics, stroke, clinical decision making, neuropsychology

INTRODUCTION

Egocentric spatial neglect is characterised by an inability to respond to stimuli presented in the contralesional hemifield (Heilman, Bowers, Valenstein, & Watson, 1987). Neglect is commonly diagnosed using a test battery including cancellation tasks (Albert, 1973; Azouvi et al., 2006; Donnelly et al., 1999; Rorden & Karnath, 2010; Vaes et al., 2015; Weintraub & Mesulam, 1988). In these tasks, patients with egocentric neglect cancel fewer targets on the contralesional than on the ipsilesional side of space, while healthy controls cancel a high number of targets evenly throughout the search array (Dalmaijer, Stigchel, Nijboer, Cornelissen, & Husain, 2014; Donnelly et al., 1999; Gauthier, Dehaut, & Joanette, 1989).

Performance on cancellation tasks can be quantified by comparing the number of cancelled targets on the left and right side of space (R-L score) or with the Centre of Cancellation (CoC), which represents the average location of cancelled targets (e.g. Azouvi et al., 2006; Demeyere & Gillebert, 2019; Demeyere, Riddoch, Slavkova, Bickerton, & Humphreys, 2015; Robertson et al., 1994; Rorden & Karnath, 2010; Vaes et al., 2015). These measures aim to capture the *spatial asymmetry* typical of neglect, that is, the difference in the probability to cancel targets between the left and right side of the cancellation display (Huygelier & Gillebert, 2018). To establish whether a patient has neglect, the R-L or CoC score is compared to a single impairment threshold, not considering the total number of cancelled targets (e.g. Demeyere et al., 2015; Rorden & Karnath, 2010). We refer to this type of thresholds as *fixed normative cutoffs*. Fixed normative cutoffs are based on percentiles of test scores obtained in neurologically healthy individuals or stroke patients without egocentric neglect (e.g. Demeyere, et al., 2015; Rorden & Karnath, 2010). These percentiles represent the most extreme test scores that

*Correspondence and reprint requests to: Céline R. Gillebert, Department of Brain and Cognition, KU Leuven, Tiensestraat 102, Box 3711, Leuven 3000, Belgium. E-mail: celine.gillebert@kuleuven.be

can be observed in individuals without neglect and are expected to limit false-positive diagnosis.

Measurement Precision of Cancellation Performance Depends on Non-Spatial Impairments

Despite the popularity of cancellation tasks, few studies investigated their *measurement precision* (Bailey, Riddoch, & Crome, 2004; Machner, Mah, Gorgoraptis, & Husain, 2012). Measurement precision is important as it affects the certainty with which conclusions can be drawn based on test scores (Crawford & Garthwaite, 2002; Lord, 1952; Lord, Novick, & Birnbaum, 1968; Slick, 2006). In classic test theory, it is assumed that the *observed test score* of a person on a specific moment can be divided in two parts: the *true underlying test score*[1] that remains stable across test moments and the *measurement error* that leads to probabilistic variation of observed test scores across moments (Novick, 1966). Thus, the observed score of an individual is considered an estimate of the true underlying score (Lord, Novick, & Birnbaum, 1968; Slick, 2006), and *measurement precision* refers to the extent to which an observed score reflects the true score. When measurement precision is low, the chance is low that we can observe the same score across repeated assessments.

Bailey et al. (2004) and Machner et al. (2012) estimated the measurement error of the total number of cancellations and the R-L score divided by the total number of cancellations by studying the variability in these scores across repeated assessments in neglect patients. The authors observed considerable variability in cancellation scores across repeated testing. However, by estimating the measurement error across all patients, they treated the error as a constant property, reflecting the assumption that each patient's observed score is associated with the same level of error. Measurement error is indeed considered a constant property of scores in classic test theory (Novick, 1966), but this principle may not apply to cancellation scores. That is, in contrast to continuous outcome measures, responses on a cancellation task are discrete: a patient either cancels or omits a single target.

Given their discrete nature, cancellation responses are best described by a binomial distribution with the mean equal to NP and variance equal to $NP(1 - P)$, where N is the number of targets and P is the probability to cancel a target. The association between the mean and variance of the distribution (Brown, Thomas, & Patt, 2017; Lord, 1952; Lord et al., 1968; McDonald, 2011) implies that the measurement error of cancellation test scores is variable. Applying the binomial formula to a cancellation task with 50 targets reveals that we can expect, for instance, a variance of 12.5 cancelled targets across repeated tests for a patient with 50% chance to cancel each target and a variance of 4.5 cancelled targets for a patient

with a 10% or 90% chance to cancel each target. In other words, the measurement error of the number of cancelled targets is highest when the probability to cancel a target equals 50% and decreases as the probability approximates 0% or 100%. Binomial variance affects test scores that rely on the number of cancelled targets such as the R-L score in the same way. The CoC measurement error has already been shown to be variable, with the CoC standard deviation increasing as the number of cancelled targets decreases (Toraldo, Romaniello, & Sommaruga, 2017).

This non-constant error variance may be important since stroke patients likely vary in their ability to cancel targets in a non-lateralised way, or in other words, in the number of omissions made across the visual field. We will refer to the reduced probability of cancelling targets irrespective of their location in the visual field as *non-spatial errors* (Huygelier & Gillebert, 2018). Non-spatial errors can result from a mix of multiple *non-spatial impairments*, including, among others, impairments in selective and sustained attention (Foldi, Jutagir, Davidoff, & Gould, 1992) and working memory (Husain et al., 2001). The measurement error of cancellation test scores will be higher in stroke patients with versus without such non-spatial impairments. Consequently, the most extreme R-L and CoC scores that can be observed in individuals without neglect will depend on non-spatial impairments. Therefore, it is unclear whether fixed normative cutoffs can successfully limit the rate of false-positive neglect diagnoses as these cutoffs ignore non-spatial errors.

The chance of false-positive neglect diagnosis also depends on how multiple test instruments inform the diagnosis. Although it is not uncommon to diagnose neglect on a single cancellation task (e.g. Brink, Verwer, Biesbroek, Visser-Meily, & Nijboer, 2017; Demeyere & Gillebert, 2019; Farnè et al., 2004; Nijboer, Kollen, & Kwakkel, 2013), many researchers and clinicians use multiple tasks (e.g. cancellation, line bisection, and figure copying) to diagnose egocentric neglect. There is however no gold-standard approach for reconciling conflicting diagnoses from multiple tests, as diagnostic criteria vary considerably across published studies. Some studies required unanimous diagnostic agreement from two cancellation tasks (e.g. Smania et al., 1998), while others required a certain proportion of agreement across multiple tasks varying from 17% to 67% (e.g. Cazzoli et al., 2012; Dalmaijer et al., 2018; McIntosh, Schindler, Birchall, & Milner, 2005; Plummer, Morris, & Dunai, 2003; Rengachary, He, Shulman, & Corbetta, 2011; Rorden & Karnath, 2010; Urbanski et al., 2010; Verdon, Schwartz, Lovblad, Hauert, & Vuilleumier, 2010). The impact of these different methods on the rate of false positives has not been investigated yet.

The Present Study

Previous studies developed fixed cutoffs to interpret R-L and CoC scores to limit false positives below a specific threshold (e.g., Demeyere et al., 2015; Rorden & Karnath, 2010). We

---

[1]We deliberately use the term 'true underlying test score' and not 'true deficit'. Whether the true underlying test score represents the true underlying deficit depends on the validity of the task and is an independent issue from the measurement error.

investigated whether fixed cutoffs can indeed control false positives. We did not study how cutoffs affect the balance between false positives and false negatives, since their optimal balance depends on the diagnostic context (Habibzadeh, Habibzadeh, & Yadollahie, 2016). We simulated cancellation data using a simple probabilistic model (Huygelier & Gillebert, 2018). Our model merely represents non-spatial errors on cancellation tasks and makes no assumptions of the underlying non-spatial impairments or how search strategy affects performance. We used Monte Carlo simulation, a method to make inferences using random numbers that follow a certain probability distribution (i.e. in our case the binomial distribution) (Beisbart & Norton, 2012). This method allows full control over the true test score underlying the observed test score, making it valuable for psychometric research (Feinberg & Rubright, 2016) and providing theoretical insights that can aid clinical decision making (Beaujean, 2018).

Using these simulated data, we assessed the impact of *fixed cutoffs* on false-positive diagnosis. We predicted that the measurement precision of R-L scores would be lowest in simulated cases who had 50% non-spatial errors, which in turn would result in an inflated false-positive neglect diagnosis. For the CoC, we predicted the lowest measurement precision and highest false-positive rates for the highest percentages of non-spatial errors, in line with earlier findings (Toraldo et al., 2017). This procedure was compared to a new approach, in which cutoffs were *adjusted* according to the total cancelled targets. We predicted better control over false positives using adjusted cutoffs compared to fixed cutoffs. Moreover, we assessed the impact of different diagnostic methods on false positives. Neglect was either diagnosed on a single test or on multiple tests. In case of multiple tests, we compared an approach of diagnosing neglect on unanimous versus proportional agreement of test results.

Finally, we aimed to illustrate the real-world impact of our premises using cancellation data acquired from a cohort of 651 stroke patients. In this analysis, we first assessed the occurrence and the distribution of non-spatial errors in our sample of stroke patients. We then assessed the percentage of patients for which fixed normative versus adjusted cutoffs would lead to a different diagnosis. Finally, we assessed whether fixed normative versus adjusted cutoffs differed most in rates of diagnosing neglect for patients who made approximately 50% non-spatial errors.

METHOD

The Theoretical Impact of Diagnostic Methods on False-Positive Rates

Simulating cancellation data using a binomial model

To simulate cancellation data, we generated 50 uniformly distributed target locations ranging from −1 (left border of search matrix) to +1 (right border of search matrix). We chose 50 targets, since this approximates the number of targets of many cancellation tasks: the Bells test with 35 targets (Gauthier et al., 1989), line cancellation with 40 targets (Albert, 1973), diamond cancellation with 48 targets (Vaes et al., 2015), Oxford Cognitive Screen (OCS) cancellation with 50 targets (Demeyere et al., 2015), Star cancellation with 54 targets (Halligan, Cockburn, & Wilson, 1991) and the letter cancellation with 60 targets (Weintraub & Mesulam, 1985).

Then, cancellation responses for each target were simulated according to a simple binomial model (see Huygelier & Gillebert, 2018). The model assumes that cancellation responses result from a probabilistic process in which each target has a certain probability to be cancelled. In case of egocentric neglect, the probability underlying cancellation responses on the contralesional side is smaller than the probability underlying cancellation responses on the ipsilesional side of the cancellation page. The difference between these two probabilities is referred to as '*spatial asymmetry*'. For patients with no neglect, these probabilities are equal. The cancelled targets depend on these two probabilities and are randomly generated using a binomial distribution.

Importantly, our model assumes that the probability to cancel targets across the entire display can be smaller than 1, allowing for *non-spatial errors* to occur. Note that in case of no true spatial asymmetry, non-spatial errors will be directly related to overall cancellation performance. That is, if a patient has a probability of cancelling 50% of all targets across the cancellation array, then the expected total cancelled targets will be equal to 50%. In case of a true spatial asymmetry, the overall cancellation performance can reflect a combination of non-spatial and spatial errors. Using this model, cancellation responses were simulated for patients with *no* spatial asymmetry and with varying levels of non-spatial errors. The level of non-spatial errors ranged from 0 to 1 in 10 steps with 10,000 simulations for each level, producing a dataset with 110,000 simulated observations.

Finally, the R-L and CoC scores were calculated for each of these simulated cancellation tasks. The R-L score was calculated by subtracting the proportion cancelled targets for the 25 targets located on the left side from that of the 25 targets located on the right side. The CoC was calculated by averaging the location of cancelled targets and subtracting the average location of all targets. Both R-L and CoC scores ranged from −1 to +1, where negative values indicated more cancelled targets on the left than right side and vice versa.

Estimating fixed and adjusted cutoffs

The 5th and 95th percentiles of the R-L and CoC scores were estimated based on our simulated dataset. Note that, based on the law of large numbers, the percentile cutoffs will represent the true percentiles. We chose the 5th and 95th percentiles in line with the cutoffs reported by Demeyere et al. (2015). Two types of R-L and CoC cutoffs were calculated; fixed and adjusted cutoffs. To estimate the fixed cutoffs, a subset of simulated data in which at least 80% targets had been cancelled was chosen as this performance is similar to that

Table 1. Demographic and stroke characteristics of stroke sample

Variable	N^a	M	SD	Min	Max
Age	605	74	14	18	97
Years of education	375	12	3	0	23
Handedness (right/left/ ambidextrous)	323/25/5				
Gender (female/male)	291/358				
Stroke etiology (ischemia, haemorrhage, transient ischemic attack)	292/74/17				
Lesion side (right/left/bilateral)	268/237/22				
Stroke to test interval (days)	284	6	7	0	90

[a] Missing numbers represent patients whose stroke characteristics were unreported in their medical notes.

of neurologically healthy individuals as reported for instance in Demeyere et al. (2015). The fixed cutoffs are based on different expected total performances. Adjusted cutoffs were determined on subsets of simulated data, each with a specific expected proportion of cancelled targets ranging from 0 to 1 in 10 steps.

False positives for different diagnostic methods

Fixed and adjusted cutoffs were used to interpret the simulated cancellation test scores. Each of the 110,000 simulated individual performances was classified into one of three categories. For a single test administration, observations with an R-L or CoC score greater than the 95th percentile were classified as left neglect, while observations with an R-L score lower than the 5th percentile were classified as right neglect. Single observations with R-L or CoC scores between the 5th and 95th percentile were classified as no neglect. This classification was performed separately for the fixed and adjusted cutoffs. For the unanimous agreement, the number of administered tests varied from 2 to 3 and neglect was only diagnosed when each of the administered tests consistently indicated neglect for the same side of space (i.e. consistent left-sided or consistent right-sided neglect). For the proportional agreement, we simulated 5 test administrations and varied the minimum number of positive test results required for diagnosis from 1 to 4. Observed scores across simulated retests varied probabilistically, while the true scores remained constant. Then, for each classification method, we calculated the rate of false-positive diagnoses. Note that false positives should remain below 10% when testing for the presence of left- and right-sided neglect (two-sided testing) when using the 5th and 95th percentiles and a single test. That is, as 10% of the control group without neglect have test scores that exceed these percentiles, you expect 10% false-positive diagnosis. We checked whether false positives indeed remained below the 10% threshold, as this threshold has been used in previous studies (e.g., Demeyere et al., 2015). For unanimous agreement, false positives should be 1% and

0.1% for 2 and 3 unanimous tests respectively. For proportional agreement, false positives should follow $1 - [(1 - Q^k)^m]$, with Q the probability of a false positive, k the number of false positives and m the number of tests. Note that false positives refer to simulated cases that were classified as neglect patients of which the true R-L and CoC score was zero. However, in case of a test that is insufficiently able to detect neglect, a true test score of zero does not necessarily proof the absence of neglect.

The Real-World Impact of Fixed Normative versus Adjusted Cutoffs on Neglect Diagnosis

A consecutive sample of stroke survivors was recruited from the John Radcliffe Hospital (Oxford, UK) between February 2012 and September 2018 in compliance with the regulations of the National Research Ethics Service (11/WM/0299 and 14/LO/0648) and Helsinki declaration. Patients were included if they were able to remain alert for 20 min and were able to provide informed consent. Participant characteristics are reported in Table 1.

The OCS cancellation task is a search matrix of 150 heart drawings pseudo-randomly scattered across a landscape orientation A4 page. Two-thirds of these drawings have left or right gaps (distractors) and the remaining third are complete drawings (targets). These drawings are arranged according to a grid pattern ensuring that there are an equal number of targets and distractors across different areas of the page. Patients were asked to cross out all complete hearts. Patients were given two practice trials before proceeding to the full task. Patients who were unable to hold a pen responded by pointing to each stimulus which was then immediately marked by the examiner. Each patient was allowed 3 min to complete the task. Performance was summarised by subtracting the number of cancelled targets on the right side from the left side of this array and was divided by 20 targets. This R-L score[2] ranged from −1 to 1 and discards the central 10 targets. Demeyere et al. (2015) reported fixed normative cutoffs (5th and 95th percentiles) of −.10 and .15.

Using these empirical data, we evaluated the implications of our simulations. First, we determined the prevalence of non-spatial errors in stroke patients without statistically significant neglect according to the fixed normative cutoffs. For this purpose, we calculated the proportions of patients obtaining a proportion of cancelled targets of either 0–.20, .22–.40, .42–.60, .62–.80 and .82–1.00 cancelled targets with an R-L score $\geq -.10$ and $\leq .15$. Second, we assessed whether fixed normative versus adjusted cutoffs lead to different diagnostic decisions. To this end, the proportion of patients classified as having statistically significant neglect based on the fixed normative cutoffs (i.e. −.10 and .15) and our adjusted cutoffs was calculated. For the adjusted cutoffs, stroke patients were categorised according to their total performance in

---

[2]These analyses focus on the R-L score because the data did not contain information of each single target location which impeded the computation of CoC scores.

Table 2. Fixed and adjusted cutoffs for the two measures of cancellation performance

	CoC		R-L	
	5th Pc	95th Pc	5th Pc	95th Pc
Fixed cutoffs				
$p* \geq .80$	−.05	.05	−.12	.12
Adjusted cutoffs				
$p = .00$	.00	.00	.00	.00
$p = .10$	−.51	.51	−.12	.12
$p = .20$	−.33	.33	−.20	.20
$p = .30$	−.24	.24	−.20	.20
$p = .40$	−.19	.19	−.24	.24
$p = .50$	−.15	.15	−.24	.24
$p = .60$	−.12	.12	−.24	.24
$p = .70$	−.10	.10	−.20	.20
$p = .80$	−.08	.08	−.20	.20
$p = .90$	−.05	.05	−.12	.12
$p = 1.00$	.00	.00	.00	.00

Note: p = expected total performance. $p*$ = observed total performance. CoC = the standardised average location of all cancelled targets, R-L = the difference between the proportion of cancelled targets on the left versus right visual field. Pc = percentile.

different groups: 0–.20, .22–.40, .42–.60, .62–.80, .82–1.00 cancelled targets. Third, we assessed whether diagnostic decisions based on these two cutoffs differed most for patients who cancelled around .50 (range: .40 to .60 targets), for whom we expected the highest false-positive rates.

RESULTS

Fixed and Adjusted Cutoffs

First, fixed cutoffs were calculated based on the simulated dataset. In simulated cases where at least 80% of targets were cancelled, R-L scores more extreme than ±.12 represented statistically significant neglect impairment at the 10% level (Table 2). These cutoffs align well with fixed normative cutoffs with Demeyere et al. (2015) reporting cutoffs of −.10 and .15 and Robertson et al. (1994) reporting ±.10. CoC scores more extreme than ±.05 represented statistically significant neglect impairment at the 10% level based on our simulated data. These CoC cutoffs align well with the normative cutoffs reported by Rorden and Karnath (2010) of ±.08.

Next, adjusted cutoffs were calculated for each performance level. Adjusted R-L cutoffs ranged from ±.24 in cases where 50% of targets were expected to be cancelled to ±.12 in cases where 90% or 10% of targets were expected to be cancelled (Table 2). Adjusted CoC cutoffs ranged from ±.51 in cases where 10% of targets were expected to be cancelled to ±.05 in cases where 90% of targets were expected to be cancelled. These adjusted cutoffs illustrate that extreme R-L scores were most likely when the probability to cancel targets was between 40% and 60%. Extreme CoC scores were most likely when the probability to cancel targets was equal to 10%.

False-Positive Rates for Different Diagnostic Methods

A single test administration

The rate of false positives was calculated based on fixed and adjusted cutoffs for a single test administration (Figure 1A). When R-L scores and fixed cutoffs were used to diagnose neglect, the rate of false positives increased as the non-spatial errors approximated 40–60%. False positives reached a maximum of 30% when the non-spatial errors equalled 50%. False positives exceeded the 10% threshold for non-spatial errors ranging from 20% to 80%. However, when using adjusted cutoffs, false positives remained below the 10% threshold for non-spatial errors between 20% and 80%. When CoC scores and fixed cutoffs were used to diagnose neglect, false positives increased as the non-spatial errors increased. The false positives exceeded the 10% threshold for all non-spatial errors larger than 10% and reached 90% for the highest level of non-spatial errors. The adjusted CoC cutoffs produced better control over false positives as they remained equal to or below the 10% threshold for non-spatial errors from 20% to 80%.

Multiple tests – unanimous agreement

Next, the false-positive rate was calculated in simulated cases where multiple consistent test results were required to diagnose neglect. In simulated cases where unanimous agreement across multiple tests was required for diagnosis, false positives decreased as the number of tests increased both when using fixed and adjusted cutoffs (Figures 1B and 1C). When using fixed cutoffs, R-L false positives dropped below the 10% level when two consistent positive test results were required (Figure 1B). CoC false positives dropped below the 10% threshold when three consistent positive test results were required (Figure 1C). False positives for the adjusted cutoffs were below 1% and 0.1% for 2 and 3 unanimous tests, respectively.

Multiple tests – proportional agreement

Finally, the rate of false positives was calculated for simulated cases in which positive test results from only a proportion of administered tests was required to diagnose neglect. In simulated cases where 1/5 positive test results were required for diagnosis, false positives exceeded the 10% threshold for almost all cases of non-spatial errors when using fixed cutoffs (Figure 2A) and adjusted cutoffs (Figure 2B) for both R-L and CoC scores. More specifically, for fixed cutoffs false positives ranged from 40% to 80% for R-L scores and from 40% to 100% for CoC scores for non-spatial errors ranging from 10% to 90%. Thus, false positives exceeded the expected 10% rate by a fourfold to a ninefold. When fixed cutoffs were used, false positives fell below the 10% threshold for R-L scores when 3/5 positive tests were required, but remained above this threshold for CoC scores even when 4/5
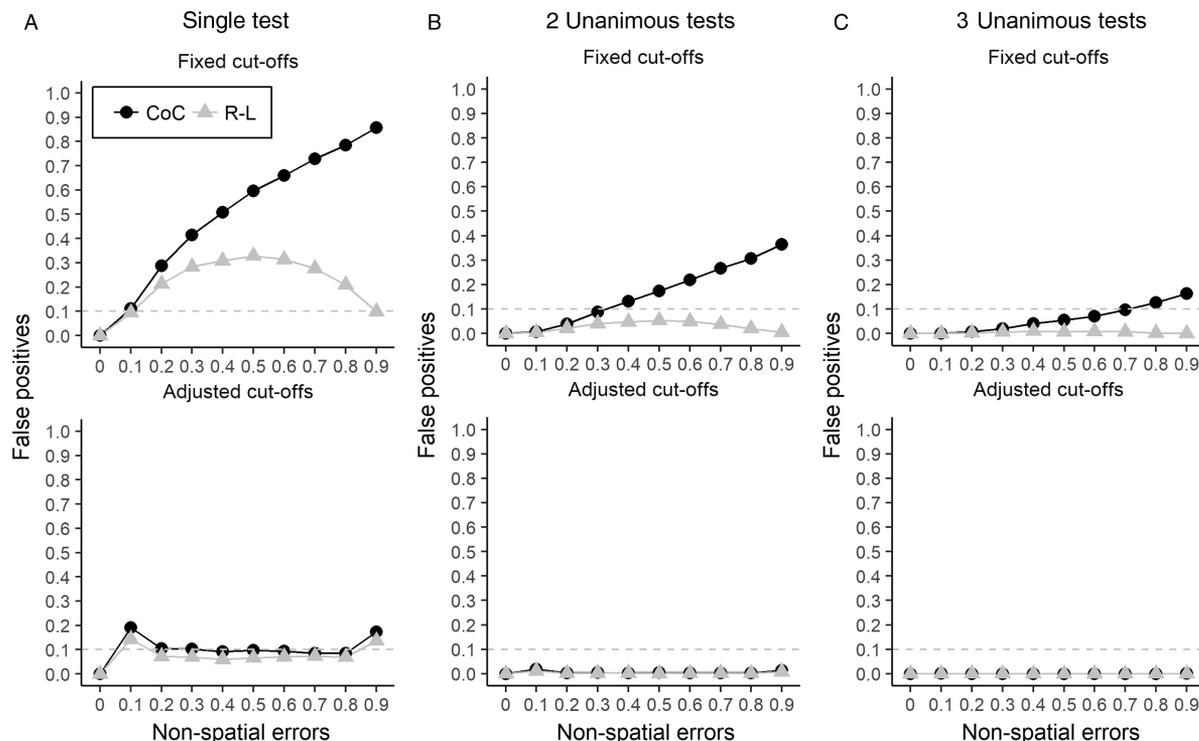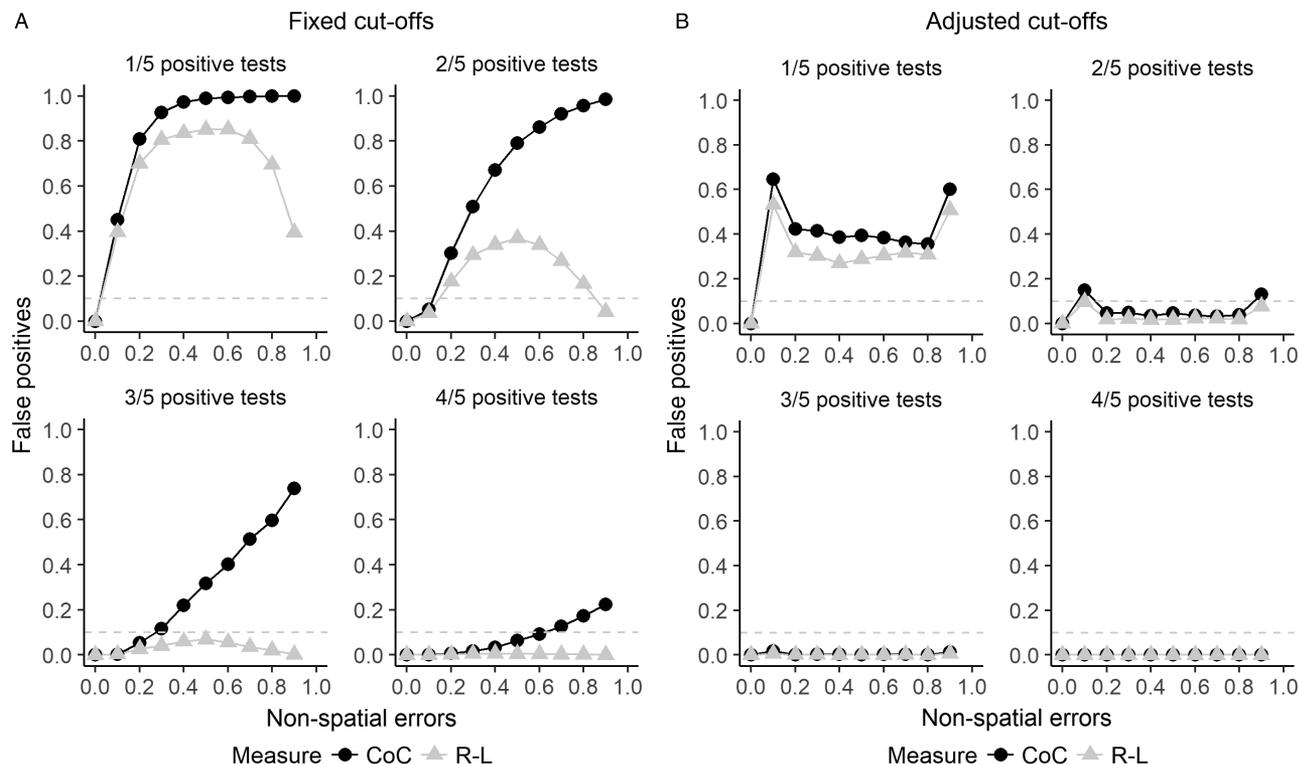
Fig. 1. The simulated false positives as a function of the expected non-spatial errors for the CoC and R-L score for fixed cutoffs and adjusted cutoffs and for basing the diagnosis on a single test result (A) or on unanimous positive tests results based on two tests (B) and three tests (C). The dashed grey line represents the maximum expected false-positive rate based on the 5th and 95th percentiles (10%). The CoC is the standardised average location of all cancelled targets. The R-L is the difference between the proportion of hits on the left and right side of the cancellation array.

Fig. 2. The simulated false positives as a function of the expected non-spatial errors for the CoC and R-L score for fixed cutoffs (A) and adjusted cutoffs (B) and for basing the diagnosis on a minimum of one, two, three, four positive tests out of five administered tests. The dashed grey line represents the maximum expected false-positive rate based on the 5th and 95th percentiles (10%). The CoC is the standardised average location of all cancelled targets. The R-L is the difference between the proportion of hits on the left and right side of the cancellation array.
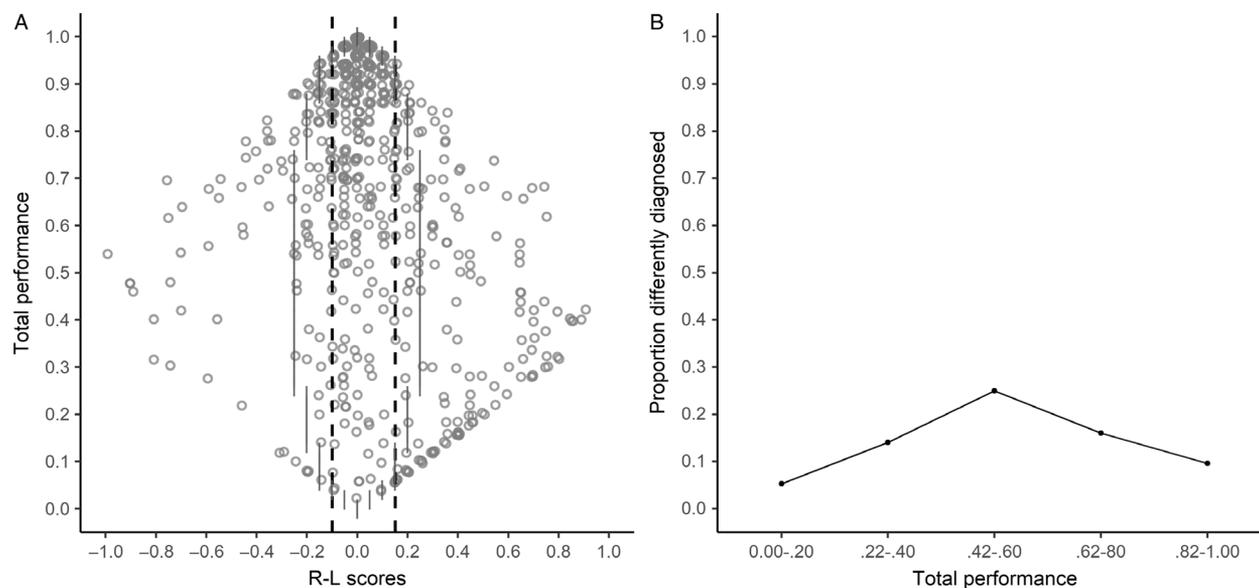
Fig. 3. The relation between the total performance on the hearts cancellation task and R-L scores. The relation of R-L scores and total performance on the hearts cancellation task of 655 stroke patients (A). The grey circles represent data of the stroke patients. The black dashed line represents the fixed cutoffs −.10 and .15. The dark grey lines represent adjusted cutoffs based on the binomial model. The proportion of patients classified as a neglect patient according to the fixed cutoffs and not according to the adjusted cutoffs is visualised as a function of the total performance on the hearts cancellation task (B).

positive tests were required (Figure 2A). Alternatively, when adjusted cutoffs were used, false positives dropped to the 10% level for R-L and CoC cutoffs when at least 2/5 positive results were required to diagnose neglect (Figure 2B). For adjusted cutoffs, false positives approximated the expected rates of 41%, 5%, 0.5%, and 0.005% for 1/5, 2/5, 3/5, 4/5 positive results, respectively.

The Real-World Impact of Fixed Normative versus Adjusted Cutoffs on Neglect Diagnosis

First, we estimated the prevalence of non-spatial errors in stroke patients without spatial asymmetry in cancellation performance. These patients were identified by comparing cancellation responses to the fixed normative cutoffs published by Demeyere et al. (2015). The results showed that 394 out of 651 stroke patients did not obtain an R-L score indicative of egocentric neglect. Figure 3A illustrates that a considerable number of these patients cancelled a low number of targets. More specifically, 28 of these 394 patients with normal R-L scores cancelled between 0% and 20% targets, 22 patients cancelled between 22% and 40% targets, 33 patients cancelled between 42% and 60% of targets, 69 patients cancelled between 62% and 80% of targets and 242 patients cancelled more than 80% of targets. Thus, 119 of these 394 patients who were not diagnosed with neglect according to fixed normative cutoffs (i.e. no spatial asymmetry in cancellation performance) cancelled between 20% and 80% of all targets. It is likely that their cancellation performance was affected by non-spatial impairments as they did not show a large difference in cancelled targets between

both sides of the visual field, but still failed to cancel many targets. This group of patients represents 19% of our entire stroke cohort. These results suggest that the inflated false-positive rates in our simulations that applied to simulated cases with non-spatial errors from 20% to 80% and without spatial asymmetries could apply to 19% of all stroke patients[3].

Patient cancellation responses were also compared to the fixed normative and the new adjusted cutoffs. As previously mentioned, when compared to fixed normative cutoffs, 394 patients did not show statistically significant neglect, implying that 257 patients exhibited statistically significant egocentric neglect. Of these 257 patients, 85 were not diagnosed with neglect according to the adjusted cutoffs. All patients who were diagnosed as neglect patients using the adjusted cutoffs were also diagnosed as neglect patients using the fixed normative cutoffs. The discrepancy between the number of patients diagnosed according to fixed normative versus adjusted cutoffs was highest for patients who cancelled between 40% and 60% of targets (Figure 3B).

DISCUSSION

While fixed normative cutoffs of cancellation test scores are frequently used for diagnosing egocentric neglect, our study shows that they do not adequately consider the discrete nature of cancellation responses. Discrete responses are best modeled by the binomial distribution of which the variance is

---

[3]This may represent an underestimation of the percentage of patients showing no neglect and only non-spatial impairments, as the fixed normative cutoffs can produce false-positive diagnosis and we only estimated the prevalence of non-spatial impairments on those patients who were not diagnosed as neglect patients using these liberal cutoffs.

not constant. This non-constant error variance implies that more extreme R-L scores are more likely to occur when the probability to make responses equals 50% across the cancellation task. By simulating data using a binomial model, we showed that extreme R-L scores were indeed most probable when 50% of targets were expected to be cancelled. In contrast, extreme CoC scores were more likely to occur when fewer targets were expected to be cancelled. This pattern of CoC error variance can be explained by the fact that the CoC relies on averaging the location of cancelled targets and the statistical average is less precise when it is based on a small sample (i.e. number of cancelled targets).

Our findings on the inflated false-positive rate for the CoC were consistent with an earlier study using a logistic model of spatial neglect (Toraldo et al., 2017). Toraldo et al. (2017) predicted inflated false-positive rates as high as 97% when the number of cancelled targets decreases to a minimum and provided a standardised CoC score to correct for this inflation of false positives associated to non-spatial impairments. Our study adds onto these earlier findings by revealing the impact of multiple tests on false positives, by providing a direct comparison of false-positive rates for the CoC and R-L scores and by showing the impact of the non-constant error variance on real-world neglect diagnosis.

Patients with Non-spatial Impairments More Likely Misdiagnosed as Neglect Patients

Our simulations showed that false positives depended on non-spatial errors when diagnosis was based on fixed cutoffs. This suggests that fixed cutoffs lead to more liberal diagnostic decisions for certain patients just because those patients cancel fewer targets overall. In other words, our results suggest that patients with non-spatial impairments are more likely misdiagnosed as neglect patients. Real-world data of a large stroke cohort showed that a considerable number of patients likely have non-spatial impairments. Specifically, 119 of the 394 patients without neglect showed signs of non-spatial impairments in our stroke cohort. Moreover, we showed that fixed normative versus adjusted cutoffs led to a difference in the estimated prevalence of neglect of 13%. These findings suggest that inflated false positives associated with non-spatial impairments have impact on the diagnostic specificity for a considerable number of stroke patients. As such, our analyses showed that accurate theory about the measurement precision of cancellation scores has implications for clinical practice. Neglect patients also suffer from non-spatial impairments in working memory, selective and sustained attention, among others, and these impairments may even interact with spatial impairments (Husain et al., 2001; Husain & Rorden, 2003; Husain, Shapiro, Martin, & Kennard, 1997; Robertson, 2001; Robertson, Mattingley, Rorden, & Driver, 1998; Robertson, Tegnér, Tham, Lo, & Nimmo-smith, 1995). Therefore, one may wonder why it matters that non-spatial impairments impact the chance of diagnosing a spatial impairment. We argue that it is important to measure

constructs independently in order to investigate their interaction. That is, we would not be able to establish the true association of spatial and non-spatial impairments, if the measurements of both impairments itself are dependent.

Adjusted Cutoffs and Unanimous Test Results Control False Positives

The results of this study provide two concrete ways to limit false positives to the 10% threshold[4]. First, adjusted cutoffs produced a stable rate of false positives that did not depend on the non-spatial errors. Thus, applying adjusted cutoffs leads to diagnostic decisions that are equally liberal for all patients, independent of whether these patients have non-spatial impairments. Second, requiring unanimous positive test results drastically reduced the rate of false positives, even for fixed cutoffs. For the R-L fixed cutoffs, requiring unanimous agreement of two tests effectively reduced false positives and for CoC fixed cutoffs unanimous agreement of three tests effectively reduced false positives. Some may argue that the ease of interpreting test scores using a fixed cutoff outweighs the benefit of reducing false positives. To illustrate the ease of using adjusted cutoffs, we added an explanation and simulated adjusted cutoffs that can be used to interpret R-L scores of the OCS cancellation task in Supplementary Material 1. Our adjusted cutoffs for R-L scores are equivalent to the z-test of proportions (Newcombe, 1998). Based on this statistical test implemented in our online application (http://www.psytests.be/stats/cancellation_task), one can obtain interval estimates of R-L scores for patients.

Multiple Tests Inflate False Positives

Many researchers and clinicians already integrate data of multiple tests to diagnose neglect. However, rather than requiring unanimous agreement of multiple tests, it is not uncommon to use a single positive test result out of multiple tests as the criterion to diagnose neglect (e.g. McIntosh et al., 2005; Rengachary et al., 2011; Verdon et al., 2010). The results of our simulations reveal that this method dramatically increased false positives when both fixed and adjusted cutoffs were used. This result is not surprising given that the chance that at least one test result will be a false positive increases as the number of tests administered increases (Farcomeni, 2008; Nichols & Hayasaka, 2003). The true test scores did not vary in our simulated retests, which differs from clinical practice where information across multiple test instruments that differ in construct validity (e.g., line bisection, cancellation, and figure copying tasks) is integrated. This raises the question to what extent our estimated false positives for multiple tests generalize to clinical practice. Indeed, for patients who have spatial impairments, performance on different tests (e.g. line

---

[4]Our results can be extrapolated to other levels of false positives. The same methods can be used to limit false positives to each threshold that a researcher or clinician aims to achieve by using different percentiles.

bisection, cancellation, and figure copying tasks) is associated (e.g. Azouvi et al., 2003; Bailey et al., 2004; Rorden & Karnath, 2010; Sperber & Karnath, 2016), and combining tests may reduce the chance of a misdiagnosis. However, for patients without a spatial deficit, multiple tests act as independent tests. The association between multiple tests for patients without true spatial deficits does not improve if tests are combined that differ in construct validity. Thus, although combining different tests can increase the chance of detecting a true neglect case, it will simultaneously increase false positives if no correction for multiple comparisons is applied.

Reducing False Positives Does Not Necessarily Improve Diagnostic Accuracy

Throughout this investigation, we have focused on false-positive diagnosis. However, accurate diagnosis requires a minimal rate of false positives as well as false negatives. One may wonder whether the adjusted cutoffs are not too conservative, risking to reduce false positives at the cost of detecting true neglect. Indeed, changing cutoffs will only re-weigh the balance between false positives and false negatives and it cannot simultaneously reduce both errors. However, the goal of detecting 100% of true neglect cases, cannot entirely justify inflated false positive rates as high as the rates that we found. That is, consider a scenario in which the true prevalence of neglect equals 20% and we can detect 100% of these cases, than the total percentage accurate diagnosis drops below the guess rate when false positives exceed 60%. Thus, our results signal an important issue that requires consideration.

Since we have not established how adjusted cutoffs affect the balance between false positives and false negatives, we can only recommend the use of adjusted cutoffs for contexts in which false positives need to be controlled. For instance, when including patients in a clinical trial or costly therapy, one may wish to avoid including patients who may not benefit from treatment. Moreover, we advise to use cutoffs that give equal weight to false positives and false negatives for epidemiological neglect research. Future research is necessary to develop cutoffs that balance false positives and false negatives, while considering the impact of non-spatial errors on false positives.

Search strategy and performance measures

Moreover, the current study focused on a simplified representation of clinical decision making. In clinical practice, neglect diagnosis is informed by test scores, but clinicians can integrate behavioural observations with test scores. Research has shown that the ways in which patients search targets can provide valuable information (Behrmann, Watt, Black, & Barton, 1997; Jalas, Lindell, Brunila, Tenovuo, & Hamalainen, 2002), while current practice mostly relies on performance rather than strategy measures (Dalmaijer et al., 2014). Search strategy can indeed impact performance measures.Chatterjee, Mennemeier, and Heilman (1992)

showed that instructing a patient to alternate cancelling targets on the left and right side, re-distributed cancellations, with many cancellations on both sides of the page and few cancellations in the center. Thus, future research into the clinical relevance of strategy measures relative to performance measures is necessary.

Can we still consider cancellation tasks a valid method to assess neglect?

There has been debate about which task is most valid to assess neglect. Some have argued that cancellation tasks should be preferred over line bisection tasks (Ferber & Karnath, 2001; Sperber & Karnath, 2016), but through changing the quantification of line bisection performance, McIntosh et al. found better convergence between cancellation and bisection tasks (McIntosh, Ietswaart, & Milner, 2017; McIntosh et al., 2005). Similarly, the way in which cancellation performance is quantified can affect its construct validity (Huygelier & Gillebert, 2018). In sum, more research is required to understand how each task and measure represents neglect.

CONCLUSIONS

To conclude, when you want to control false-positive neglect diagnoses, we recommend to base diagnosis on adjusted cutoffs that account for the total performance. Alternatively, if you want to use fixed cutoffs, we recommend to use the R-L score and a criterion of two unanimous test results or the CoC score and a criterion of three unanimous test results.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The researchers declare no conflicts of interest.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355617720000041

REFERENCES

Albert, M.L. (1973). A simple test of visual neglect. *Neurology*, *23*(6), 658–664. doi: 10.1212/WNL.23.6.658

Azouvi, P., Bartolomeo, P., Beis, J.-M., Perennou, D., Pradat-Diehl, P., & Rousseaux, M. (2006). A battery of tests for the quantitative assessment of unilateral neglect. *Restorative Neurology and Neuroscience*, 24(4–6), 273–285.

Azouvi, P., Olivier, S., de Montety, G., Samuel, C., Louis-Dreyfus, A., & Tesio, L. (2003). Behavioral assessment of unilateral neglect: Study of the psychometric properties of the Catherine Bergego Scale. *Archives of Physical Medicine and Rehabilitation*, 84(1), 51–57. doi: 10.1053/apmr.2003.50062

Bailey, M.J., Riddoch, M.J., & Crome, P. (2004). Test–retest stability of three tests for unilateral visual neglect in patients with stroke: Star cancellation, line bisection, and the baking tray task. *Neuropsychological Rehabilitation*, 14(4), 403–419. doi: 10.1080/09602010343000282

Beaujean, A.A. (2018). Simulating data for clinical research: A tutorial. *Journal of Psychoeducational Assessment*, 36(1), 7–20. doi: 10.1177/0734282917690302

Behrmann, M., Watt, S., Black, S.E., & Barton, J.J.S. (1997). Impaired visual search in patients with unilateral neglect: An oculographic analysis. *Neuropsychologia*, 35(11), 1445–1458. doi: 10.1016/S0028-3932(97)00058-4

Beisbart, C. & Norton, J.D. (2012). Why Monte Carlo Simulations are inferences and not experiments. *International Studies in the Philosophy of Science*, 26(4), 403–422. doi: 10.1080/02698595.2012.748497

Brink, A.F.T., Verwer, J.H., Biesbroek, J.M., Visser-Meily, J.M.A., & Nijboer, T.C.W. (2017). Differences between left- and right-sided neglect revisited: A large cohort study across multiple domains. *Journal of Clinical and Experimental Neuropsychology*, 39(7), 707–723. doi: 10.1080/13803395.2016.1262333

Brown, G.G., Thomas, M.L., & Patt, V. (2017). Parametric model measurement: Reframing traditional measurement ideas in neuropsychological practice and research. *The Clinical Neuropsychologist*, 31(6–7), 1047–1072. doi: 10.1080/13854046.2017.1334829

Cazzoli, D., Müri, R.M., Schumacher, R., von Arx, S., Chaves, S., Gutbrod, K., Bohlhalter, S., Bauer, D., Vanbellingen, T., Bertschi, M., Kipfer, S., Rosenthal, C.R., Kennard, C., Bassetti, C.L., & Nyffeler, T. (2012). Theta burst stimulation reduces disability during the activities of daily living in spatial neglect. *Brain: A Journal of Neurology*, 135(Pt 11), 3426–3439. doi: 10.1093/brain/aws182

Chatterjee, A., Mennemeier, M., & Heilman, K.M. (1992). Search patterns and neglect: A case study. *Neuropsychologia*, 30(7), 657–672. doi: 10.1016/0028-3932(92)90070-3

Crawford, J.R., & Garthwaite, P.H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, 40(8), 1196–1208. doi: 10.1016/S0028-3932(01)00224-X

Dalmaijer, E.S., Li, K.M.S., Gorgoraptis, N., Leff, A.P., Cohen, D.L., Parton, A.D., Husain, M., & Malhotra, P.A. (2018). Randomised, double-blind, placebo-controlled crossover study of single-dose guanfacine in unilateral neglect following stroke. *Journal of Neurology Neurosurgery & Psychiatry*, 89, 593–598. doi: 10.1136/jnnp-2017-317338

Dalmaijer, E.S., Stigchel, S.V. der, Nijboer, T.C.W., Cornelissen, T.H.W., & Husain, M. (2014). CancellationTools: All-in-one software for administration and analysis of cancellation tasks. *Behavior Research Methods*, 47(4), 1065–1075. doi: 10.3758/s13428-014-0522-7

Demeyere, N., & Gillebert, C.R. (2019). Ego- and allocentric visuospatial neglect: Dissociations, prevalence, and laterality in acute stroke. *Neuropsychology*, 33(4), 490–498. doi: 10.1037/neu0000527

Demeyere, N., Riddoch, M.J., Slavkova, E.D., Bickerton, W.-L., & Humphreys, G.W. (2015). The Oxford Cognitive Screen (OCS): Validation of a stroke-specific short cognitive screening tool. *Psychological Assessment*, 27(3), 883–894. doi: 10.1037/pas0000082

Donnelly, N., Guest, R., Fairhurst, M., Potter, J., Deighton, A., & Patel, M. (1999). Developing algorithms to enhance the sensitivity of cancellation tests of visuospatial neglect. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc.*, 31(4), 668–673.

Farcomeni, A. (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Statistical Methods in Medical Research*, 17(4), 347–388. doi: 10.1177/0962280206079046

Farnè, A., Buxbaum, L.J., Ferraro, M., Frassinetti, F., Whyte, J., Veramonti, T., Angeli, V., Coslett, H.B., & Làdavas, E. (2004). Patterns of spontaneous recovery of neglect and associated disorders in acute right brain-damaged patients. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(10), 1401–1410. doi: 10.1136/jnnp.2002.003095

Feinberg, R.A. & Rubright, J.D. (2016). Conducting simulation studies in psychometrics. *Educational Measurement: Issues and Practice*, 35(2), 36–49. doi: 10.1111/emip.12111

Ferber, S. & Karnath, H.-O. (2001). How to assess spatial neglect – Line bisection or cancellation tasks? *Journal of Clinical and Experimental Neuropsychology*, 23(5), 599–607. doi: 10.1076/jcen.23.5.599.1243

Foldi, N.S., Jutagir, R., Davidoff, D., & Gould, T. (1992). Selective attention skills in Alzheimer's disease: Performance on graded cancellation tests varying in density and complexity. *Journal of Gerontology*, 47(3), P146–P153. doi: 10.1093/geronj/47.3.P146

Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The bells test: A quantitative and qualitative test for visual neglect. *International Journal of Clinical Neuropsychology*, 11(2), 49–54.

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochemia Medica*, 26(3), 297–307. doi: 10.11613/BM.2016.034

Halligan, P.W., Cockburn, J., & Wilson, B.A. (1991). The behavioural assessment of visual neglect. *Neuropsychological Rehabilitation*, 1(1), 5–32. doi: 10.1080/09602019108401377

Heilman, K.M., Bowers, D., Valenstein, E., & Watson, R.T. (1987). Hemispace and hemispatial neglect. In M. Jeannerod (Ed.), *Neurophysiological and neuropsychological aspects of spatial neglect*. Advances in Psychology, Vol. 45, (pp. 115–150). Amsterdam, The Netherlands: North-Holland.

Husain, M., Mannan, S., Hodgson, T., Wojciulik, E., Driver, J., & Kennard, C. (2001). Impaired spatial working memory across saccades contributes to abnormal search in parietal neglect. *Brain*, 124(5), 941–952. doi: 10.1093/brain/124.5.941

Husain, M. & Rorden, C. (2003). Non-spatially lateralized mechanisms in hemispatial neglect. *Nature Reviews Neuroscience*, 4(1), 26–36. doi: 10.1038/nrn1005

Husain, M., Shapiro, K., Martin, J., & Kennard, C. (1997). Abnormal temporal dynamics of visual attention in spatial neglect patients. *Nature*, 385(6612), 154–156. doi: 10.1038/385154a0

Huygelier, H. & Gillebert, C.R. (2018). Quantifying egocentric spatial neglect with cancellation tasks: A theoretical validation. *Journal of Neuropsychology*. doi: 10.1111/jnp.12177

Jalas, M.J., Lindell, A.B., Brunila, T., Tenovuo, O., & Hamalainen, H. (2002). Initial rightward orienting bias in clinical tasks: Normal subjects and right hemispheric stroke patients with and without neglect. *Journal of Clinical and Experimental Neuropsychology*, 24(4), 479–490. doi: 10.1076/jcen.24.4.479.1036

Lord, F. (1952). A theory of test scores. *Psychometric Monographs*, Vol. 7 (pp. x, 84–x, 84). Richmond, VA: Psychometric Corporation.

Lord, F.M., Novick, M.R., & Birnbaum, A. (1968). *Statistical Theories of Mental Test Scores*. Oxford, England: Addison-Wesley.

Machner, B., Mah, Y.-H., Gorgoraptis, N., & Husain, M. (2012). How reliable is repeated testing for hemispatial neglect? Implications for clinical follow-up and treatment trials. *J Neurol Neurosurg Psychiatry*, 83(10), 1032–1034. doi: 10.1136/jnnp-2012-303296

McDonald, R.P. (2011). *Test theory: A unified treatment*. New York: Routledge Taylor & Francis Group.

McIntosh, R.D., Ietswaart, M., & Milner, A.D. (2017). Weight and see: Line bisection in neglect reliably measures the allocation of attention, but not the perception of length. *Neuropsychologia*, 106, 146–158. doi: 10.1016/j.neuropsychologia.2017.09.014

McIntosh, R.D., Schindler, I., Birchall, D., & Milner, A.D. (2005). Weights and measures: A new look at bisection behaviour in neglect. *Cognitive Brain Research*, 25(3), 833–850. doi: 10.1016/j.cogbrainres.2005.09.008

Newcombe, R.G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine*, 17(8), 873–890. doi: 10.1002/(SICI)1097-0258(19980430)17:8<873::AID-SIM779>3.0.CO;2-I

Nichols, T. & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research*, 12(5), 419–446. doi: 10.1191/0962280203sm341ra

Nijboer, T.C., Kollen, B.J., & Kwakkel, G. (2013). Time course of visuospatial neglect early after stroke: A longitudinal cohort study. *Cortex*, 49(8), 2021–2027.

Novick, M.R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1), 1–18. doi: 10.1016/0022-2496(66)90002-2

Plummer, P., Morris, M.E., & Dunai, J. (2003). Assessment of unilateral neglect. *Physical Therapy*, 83(8), 732–740.

Rengachary, J., He, B.J., Shulman, G., & Corbetta, M. (2011). A behavioral analysis of spatial neglect and its recovery after stroke. *Frontiers in Human Neuroscience*, 5, 29. doi: 10.3389/fnhum.2011.00029

Robertson, I.H. (2001). Do we need the "lateral" in unilateral neglect? Spatially nonselective attention deficits in unilateral neglect and their implications for rehabilitation. *Neuroimage*, 14(1), S85–S90. doi: 10.1006/nimg.2001.0838

Robertson, I.H., Halligan, P.W., Bergego, C., Hömberg, V., Pizzamiglio, L., Weber, E., & Wilson, B.A. (1994). Right neglect following right hemisphere damage? *Cortex*, 30(2), 199–213. doi: 10.1016/S0010-9452(13)80193-1

Robertson, I.H., Mattingley, J.B., Rorden, C., & Driver, J. (1998). Phasic alerting of neglect patients overcome their spatial deficit in visual awareness *Nature*, 395(6698), 169–172. doi: 10.1038/25993

Robertson, I.H., Tegnér, R., Tham, K., Lo, A., & Nimmo-smith, I. (1995). Sustained attention training for unilateral neglect: Theoretical and rehabilitation implications. *Journal of Clinical and Experimental Neuropsychology*, 17(3), 416–430. doi: 10.1080/01688639508405133

Rorden, C. & Karnath, H.-O. (2010). A simple measure of neglect severity. *Neuropsychologia*, 48(9), 2758–2763. doi: 10.1016/j.neuropsychologia.2010.04.018

Slick, D.J. (2006). Psychometrics in neuropsychological assessment. In E. Strauss, E.M. Sherman, & O. Spreen (Eds.), *A Compendium of Neuropsychological Tests: Administration, norms and commentary* (pp. 3–43). New York: Oxford University Press.

Smania, N., Martini, M.C., Gambina, G., Tomelleri, G., Palamara, A., Natale, E., & Marzi, C.A. (1998). The spatial distribution of visual attention in hemineglect and extinction patients. *Brain*, 121(9), 1759–1770. doi: 10.1093/brain/121.9.1759

Sperber, C. & Karnath, H.-O. (2016). Diagnostic validity of line bisection in the acute phase of stroke. *Neuropsychologia*, 82(Feb), 200–204. doi: 10.1016/j.neuropsychologia.2016.01.026

Toraldo, A., Romaniello, C., & Sommaruga, P. (2017). Measuring and diagnosing unilateral neglect: A standardized statistical procedure. *The Clinical Neuropsychologist*, 31(6–7), 1248–1267. doi: 10.1080/13854046.2017.1349181

Urbanski, M., Schotten, M.T. de, Rodrigo, S., Oppenheim, C., Touzé, E., Méder, J.-F., Moreau, K., Loeper-Jeny, C., Dubois, B., & Bartolomeo, P. (2010). DTI-MR tractography of white matter damage in stroke patients with neglect. *Experimental Brain Research*, 208(4), 491–505. doi: 10.1007/s00221-010-2496-8

Vaes, N., Lafosse, C., Nys, G., Schevernels, H., Dereymaeker, L., Oostra, K., Hemelsoet, D., & Vingerhoets, G. (2015). Capturing peripersonal spatial neglect: An electronic method to quantify visuospatial processes. *Behavior Research Methods*, 47(1), 27–44. doi: 10.3758/s13428-014-0448-0

Verdon, V., Schwartz, S., Lovblad, K.-O., Hauert, C.-A., & Vuilleumier, P. (2010). Neuroanatomy of hemispatial neglect and its functional components: A study using voxel-based lesion-symptom mapping. *Brain*, 133(3), 880–894. doi: 10.1093/brain/awp305

Weintraub, S. & Mesulam, M.M. (1985). Mental state assessment of young and elderly adults in behavioral neurology, In M.M. Mesulam (Ed.), *Principles of behavioral neurology*, (pp. 71–123). Philadelphia: F.A. Davis.

Weintraub, S. & Mesulam, M.M. (1988). Visual hemispatial inattention: Stimulus parameters and exploratory strategies. *Journal of Neurology, Neurosurgery, and Psychiatry*, 51(12), 1481–1488. doi: 10.1136/jnnp.51.12.1481