'WEBSITE MATERIAL'

Update: July, 4th, 2018

Measuring and diagnosing neglect: a standardized statistical procedure

Alessio Toraldo[1], Cristian Romaniello[1], Paolo Sommaruga[2]

[1]*Department of Brain and Behavioural Sciences, University of Pavia, Italy*
[2]*Reti Televisive Italiane, Technology Operative Center, Segrate (Milano), Italy*

Correspondence concerning this article should be addressed to *Alessio Toraldo*, Department of Brain and Behavioural Sciences, Università di Pavia, Piazza Botta 11, 27100 Pavia, Italy; Phone +39 0382 986131; Fax: +39 0382 986272; E-mail: alessio.toraldo@unipv.it.

Other authors' details: Cristian Romaniello: cristian.romaniello01@universitadipavia.it; Paolo Sommaruga: paolosom@gmail.com.

Abbreviations: *MPH*, Mean Position of Hits; *MPO*, Mean Position of Omissions; *MPT*, Mean Position of Targets; *MOH*, Mean Ordinal position of Hits; *MdnPH*, Median Position of Hits; *MPP*, Mean Position of Positions; *LCR-adjusted*, Left-Centre-Right-adjusted; *C-adjusted*, Centre-adjusted (or centred); *CoC*, Centre of Cancellation; *SD*, expected Standard Deviation of *MPH* (unless otherwise stated); *BF*, Bayes Factor; *H*, number of Hits; *T*, number of targets; *G*, number of target clusters; *CTL*, Central Theorem of Limits; *l*, location parameter; *s*, slope parameter; *c*, ceiling parameter; FA, False Alarm (on catch trials); CR, Correct Rejections (on catch trials); SPS, Spatial Processing Stages; EP, Extra Processing stages (different from SPS).

In this file we report many technical details that were omitted from the paper published in *The Clinical Neuropsychologist* (July 2017). In any publication, please cite:
Toraldo A, Romaniello C, Sommaruga P (2017). Measuring and diagnosing neglect: a standardized statistical procedure. *The Clinical Neuropsychologist*, 31 (6-7), 1248-1267. DOI: 10.1080/13854046.2017.1349181

The paper's Figures 1 and 2 are also included in the present text. They are numbered WM-1 and WM-2 here (WM stands for 'Website Material').

The 'Worksheet' we often refer to is the 'MPH neglect diagnosis' piece of Excel software that performs neglect diagnosis and which can be downloaded from this same Website:
psicologia.unipv.it/toraldo/mean-position-of-hits.htm.

INDEX

In the following Section we wished to clarify the logic applied when identifying an optimal index of neglect. A very basic feature of any type of measurement is a classification which attributes the same degree of the to-be-measured property to a set of measured objects; these sets which are internally homogeneous for the measured quantity are called 'equivalence classes'. Hence we wished to argue that there are indeed 'equivalence classes of neglect', i.e. families of different performances that can be stated to reflect an identical degree of neglect.

1. EQUIVALENCE CLASSES OF NEGLECT

In order to define behavioural classes sharing the same degree of neglect, we need to describe a patient's performance in a complete way. We reasoned that a full description of a patient's performance is a function relating position (in X) to Hit rate (in Y). We will imply that X is the position along the *horizontal* axis in most examples, but identical considerations hold for any other spatial dimension. There are a number of suitable equations for the XY relationship, for instance the Cumulative Gaussian distribution or logistic curves (as are those depicted in Fig. WM-1). Logistic curves have a sigmoidal shape, two asymptotes, Hit rate = 0 and 1, an inflection point at Hit rate = .5, and are characterized by two parameters, the slope s of the function and the horizontal *location* of the inflection point (l). However we need to add another parameter, the height of the upper asymptote (*ceiling*, c) – indeed while classical logistic curves have $Y=0$ and $Y=1$ as asymptotes, on a visual search task a subject might well have a top Hit rate lower than 1. Note that we are not just speaking of the part of the curve lying in the tested horizontal interval (the display) – the logistic curve has an X domain extending from $-\infty$ to $+\infty$, thus, for instance, the height Y of a curve whose upper asymptote c equals 1 might well vary between 0 and .7 in the display. Also note that if one takes a curve and changes the *ceiling* parameter, the asymptote moves from 1 to, say, .8 and 'compresses' the curve downwards. In this way the curve becomes shallower, but, mathematically, the *slope* parameter s does not change, because s is the rate at which the drop from the *ceiling* (not from 1) to zero occurs. In other words s (which we could name '*standardized slope*') is the slope one would visually observe by de-compressing any curve so that its upper asymptote is exactly 1. Thus when we speak of 'slope' all across the paper, we refer to the s parameter, and not to the apparent slope visible in a plot.

As for the lower asymptote (*floor*), this is assumed to be zero – the implications of such an assumption will be discussed later.

If any patient's performance can be summarized by a logistic curve, it will then vary by those three parameters. Hence patients can differ from one another in terms of s, l or c. Figure WM-1 shows patients varying *only* for l (A, B, C), or only for s (D, E, F), or only for c (G, H, I) – the latter set of patients have the same *standardized slope*, s.[1]

The critical step here is to decide whether or not changes in one specific parameter correspond to changes in neglect severity. As we shall see, our view is that changes in the s and l parameters correspond to changes in neglect severity, while changes in c must be assumed *not* to reflect changes in neglect.

Our first two statements are based on the 'pre-theoretical' intuition that patients A-B-C (varying only for l) have increasing neglect severity, and that patients D-E-F (varying only for s) have decreasing neglect severity. By 'pre-theoretical' we mean that all students of neglect would agree on this – or, in other words, that all neglect theories would agree on such a classification. Thus patient A produces omissions that are much more confined to the left side than are omissions by C, and patient F shows no lateral bias – no neglect – with respect to patient D who shows marked neglect.

---

[1] Another way of explaining the meaning of the three parameters is by thinking of the curve relating space to Hit rate as a Cumulative Gaussian distribution (CG), and by considering the Original Gaussian (OG) that generated it. OG has three parameters: mean (l), standard deviation (s) and 'weight' (c). The mean can be anywhere – thus a perfect 100% performance corresponds to an OG at ∞ on either side, well out of the spatial interval tested empirically. Weight is how 'heavy' OG is, 0 to 1 – in terms of the CG, weight corresponds to the superior asymptote (c, the *ceiling*), i.e. the top Hit rate that the subject would achieve at $\pm\infty$. Considering the OG distribution, Patients A, B and C vary only for the mean (l); patients D, E and F vary only for the standard deviation (s); patients G, H and I vary only for the weight (c).

Figure WM-1 *Physical position ranging from -.5 (leftmost target) to .5 (rightmost target) is plotted against Hit rate (0-1) for nine imaginary patients. Triangles along the plots' horizontal axes show the Mean Position of Hits (MPH): black, relative to the black solid curve; white, relative to the dashed curve; grey, relative to the grey curve. MPHs are identical in patients G-H-I, so the triangles overlap. Circles along the plots' upper borders show the Mean Position of Omissions (MPO) – colour conventions as before.*

Our reasoning on parameter *c* (varying between patients G-H-I) is less straightforward, perhaps, but has a clear-cut conclusion. The *c* parameter is certainly influenced by deficits that are *different* from neglect – deficits that do not produce lateral biases (hence the name 'non-lateral' deficits) and that lower the Hit rate in an identical way all across spatial positions. If we assumed that neglect severity can also vary with *c* we would be stating that telling the effects of neglect from those of non-lateral deficits would be impossible – thus invalidating all tasks of this type as diagnostic of neglect. Hence the assumption that *c* does not reflect neglect severity in any way, is *mandatory*: anyone giving up such an assumption would be giving up all neglect tests yielding the Hit/Miss-by-position data from, that is, the vast majority of neglect tests.

So we are left with two relevant parameters, *l* and *s*. We cannot choose one of them as an ideal measure of neglect severity because it would fail to detect differences in neglect severity due to the other one. Moreover, both parameters have undesirable mathematical features (they both range from −∞ to +∞, and normal subjects would have huge scalar values for both of them). The Mean Position of Hits, *MPH*, has the lucky property of combining information from *l* and from *s* (it depends on both: Fig. WM-1, A-C, D-F) and of having an intuitive and limited range (the space of the display).

It is important to highlight here that the choice of looking at the variation of a single parameter at a time was a deliberate simplification. A complete analysis would have looked also at simultaneous variations in multiple parameters – with new equivalence classes and neglect severity orders. However (i) joint variation produces a virtually infinite number of combinations, and (ii) organizing these combinations in neglect severity orders and equivalence classes would be impossible to do pre-theoretically, that is, without reference to a specific cognitive / computational / neurophysiological model of neglect. Since we aimed at developing a method with general validity (i.e. not contingent on any specific neglect theory), we decided to limit our analysis to the simplified framework exposed above.

In the following Section we report the results of the Monte-Carlo simulations which were used to choose the best index of neglect among the measures of central tendency of the distribution of Hits across physical space.

2. WHICH MEASURE OF CENTRAL TENDENCY OF THE HIT DISTRIBUTION HAS THE BEST STATISTICAL PROPERTIES?

We ran a set of Monte Carlo simulations of Hit sample distributions in virtual subjects without neglect (i.e. with a perfectly flat function relating spatial position to Hit rate) and with various degree of non-lateral deficits (i.e. with Hit rate varying all across the 0-1 range), in tasks with 10, 20, 50, 100 or 150 targets distributed in a perfectly

equispaced fashion along the horizontal axis. Each simulation generated 10,000 samples. Sample distributions of *MPH*, *MdnPH* (Median Position of Hits) and *Mid-Range* (the midpoint between the two extreme hits) were studied according to their standard deviation (SD) and Kurtosis as an index of gaussianity. Mean and Skewness values of such distributions were not studied as they were invariably zero (an expected result, given that all virtual subjects were free of lateral bias).

T->	10	10	10	150	150	150
H	**MPH**	**MdnPH**	**Mid-Range**	**MPH**	**MdnPH**	**Mid-Range**
1	.319	.319	.319			
2	.213	.213	.213			
3	.163	.219	.154	.168	.224	.159
4	.132	.165	.119			
5	.107	.155	.091			
6	.087	.121	.070			
7	.070	.106	.053			
15				.072	.116	.041
30				.047	.078	.020
45				.036	.062	.013
60				.029	.049	.009
75				.024	.041	.007
90				.019	.033	.005
105				.016	.027	.004
120				.012	.020	.003
135				.008	.014	.002
147				.003	.006	.001

Table WM-1 *Standard deviation of sample distributions (simulated with N=10,000) of MPH, MdnPH, Mid-Range with T=10 or 150. The lower the standard deviation, the more efficient the estimator.*

T->	10	10	10	150	150	150
H	**MPH**	**MdnPH**	**Mid-Range**	**MPH**	**MdnPH**	**Mid-Range**
1	−1.222	−1.222	−1.222			
2	−.664	−.664	−.664			
3	−.457	−.931	−.168	−.394	−.858	−.146
4	−.398	−.713	.106			
5	−.337	−.840	.509			
6	−.373	−.739	.610			
7	−.460	−.931	.905			
15				−.034	−.275	1.558
30				−.059	−.129	2.496
45				.049	−.080	2.885
60				−.032	−.091	3.402
75				−.053	−.148	2.815
90				.037	−.024	3.150
105				-.008	−.079	3.309
120				−.073	−.158	4.156
135				−.044	−.165	6.513
147				−.466	−.677	27.443

Table WM-2 *Kurtosis of sample distributions (simulated with N=10,000) of MPH, MdnPH, Mid-Range with T=10 or 150. Zero expresses perfect gaussianity, negative values characterize platykurtic distributions, positive values leptokurtic distributions.*

Table WM-1 shows the standard deviations of sample distributions (*N*=10,000 each) of *MPH*, *MdnPH* and *Mid-Range*. *Mid-Range* is the most efficient estimator of central tendency, because it systematically has the lowest standard deviation; *MPH* has intermediate efficiency, and *MdnPH* is the worst in this respect.

This high-efficiency of *Mid-Range* was expected on grounds of its statistical properties: *Mid-Range* is known for being the most efficient central tendency estimator for Uniform Distributions; our case does indeed resemble a sampling from a Uniform Distribution. However, *Mid-Range* also has some serious drawbacks. The first is that its distribution can be critically removed from gaussianity. Table WM-2 shows that the Kurtosis of *Mid-Range* increases very steeply as the Hits count increases. This huge leptokurticism is due to an undesirable property of *Mid-Range*: when Hit rate is relatively high, the leftmost and rightmost Hits tend very often to be the leftmost an rightmost targets – so *Mid-Range* will necessarily be zero (the display's midpoint) in all these cases, and in spite of variation in the Hit rate anywhere else. By contrast, *MPH* and *MdnPH* have much quieter Kurtosis, with values between −.5 and 0 (when both Hit and Omission counts are at least 3). This means that the standard Gaussian distribution can safely be used to compute *p*-values if *MPH* or *MdnPH* are used, while this is definitely not the case for *Mid-Range*. The other drawback of *Mid-Range* is linked to the first, and is the following. Because *Mid-Range* just considers the leftmost Hit and the rightmost Hit, patients with mild neglect would very often be misdiagnosed as normal, namely, in all cases where they detect the leftmost target even though they miss some, even many other targets in the left half of the display. To clarify this point, we simulated 100 performances by patients with mild neglect on a 50-target detection task. In the simulation, detection probability was 50% for the leftmost target, then it gradually increased in +2.5% steps, till it reached 100% in position 21/50. By using *MPH*, neglect was correctly diagnosed in 96/100 patients, thus estimating test sensitivity to be 96%. As for *Mid-Range*, clearly about half the patients would have detected the leftmost target; these same patients, of course, would have invariably detected the rightmost target, thus fixing *Mid-Range* at perfect normality, exactly halfway across the display. Hence test sensitivity would have been *at most* 50%, and probably less. Thus, *Mid-Range* not only has a very irregular distribution shape, it also has disastrously low diagnostic sensitivity in cases of mild neglect.[2] Therefore we rejected *Mid-Range*, and preferred *MPH* over *MdnPH* because it is more efficient (Table WM-1).

In the following Section, we define and discuss a full list of the assumptions that make our statistical model for diagnosing neglect a valid mathematical tool, that is, a technique with nominal false-positive rates (5% or 2% or any other desired value).

3. LIST OF ASSUMPTIONS OF THE STATISTICAL MODEL

3.1. Assumptions concerning MPH as an estimator of true neglect severity

We conceptualized true neglect severity as the unknown mean ν of the distribution whose probability density function is the Hit rate logistic curve in the space ranging from −.5 (leftmost target) to .5 (rightmost target; examples of the curve are shown in Fig. WM-1). *MPH* is an asymptotically unbiased estimator of ν (i.e. the mean *MPH* equals ν when the number of target positions = +∞)[3] if the following assumptions hold.

The first assumption concerns the shape of the logistic curve.
(1) The logistic curve is driven by three parameters: *slope*, *location* and *upper asymptote*, or '*ceiling*' (see above). The *lower asymptote*, or '*floor*' is assumed to be zero (3-parameter curve).

Another two assumptions regard the homogeneity of the distribution of targets across the display.
(2) Targets occupy positions that are equispaced along the studied spatial dimension ('**equispacing**').
(3) Each position hosts an equal number of targets ('**ties homogeneity**').

---

[2] *Mid-Range* would indeed be the best measure if neglect had been a clear-cut deficit like visual scotomas, producing a high Hit rate in a portion of the display, and a zero Hit rate anywhere else, with a very abrupt change at the boundary between the two regions. In these conditions *Mid-Range* would work optimally, as it is known to be the best estimator of Uniform Distributions' means when the external boundaries' positions are unknown. In the case of clear-cut neglect the position of the contralesional boundary would indeed be unknown. However, neglect is seldom a clear-cut phenomenon, and graded deficit distributions are far more common.
[3] This holds, exactly, if positions are equispaced and all of them have an identical number of targets.

(4) The studied spatial dimension does not covary with any other behaviourally important variable ('**univocal interpretation**').

3.2. Assumptions concerning *MPH*'s distributional shape and variability

An asymptotically unbiased estimator of true neglect level ν is not enough for diagnosis: we also need to know the shape and the variability of its distribution under the null hypothesis – these will determine statistical significance (false positive rate) and test power (diagnostic sensitivity). The model's assumption in this respect is the following.
(5) In a given normal subject or brain-damaged patient without neglect[4], all targets have exactly the same probability of being detected, no matter their nature or position ('**isoprobability**' or '**no lateral bias**' assumption).
For instance, all the bells in the Bells test (Gauthier et al., 1989) are assumed to have equal probability of being found by a given normal subject; different normal subjects can have different probabilities of finding bells, but within a subject that probability is constant. A patient without neglect might have a markedly low probability to correctly process a bell (e.g. because of amblyopia, agnosia, low motivation, etc.) but again, such a probability would be equal for all bells.

In the following Section we describe the Monte-Carlo study on *MPH* which led to the derivation of the statistical model and Equation.

4. MONTE CARLO SIMULATIONS AND MODEL EQUATION

Given assumptions (1), (2), (3) and (5) above, how exactly will *MPH* vary across repetitions of the same test, if no neglect is present?
The sample distribution of *MPH* under the null of no neglect certainly has 0 mean and is symmetrical; however, we are not aware of any analytical formula to describe its shape (Kurtosis) and its standard deviation (σ_{MPH}). To understand how Kurtosis and σ_{MPH} behave, we ran large sets of simulations. We varied the overall number of targets, *T*, in a range that covers most experimental and clinical tests: *T*=10, 20, 50, 100, 150, and the absolute number of hits, *H* = 1, 2, 3, .1*T*, .2*T*, .3*T*, .4*T*, .5*T*, .6*T*, .7*T*, .8*T*, .9*T*, *T*−3, *T*−2, *T*−1. Variations in Hit rate are contained in this set, as Hit rate = *H*/*T*. We obtained 10,000 samples for each combination.
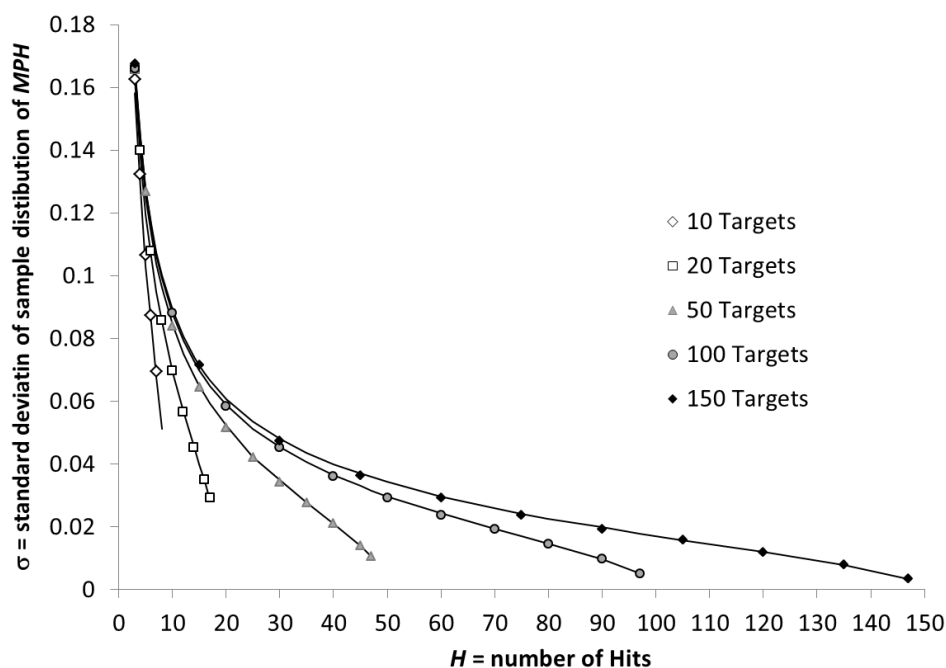
Figure WM-2 *Standard deviation of the sample MPH distribution as a function of number of targets (different plots) and Hit count (horizontal axis). Dots show values obtained from simulations of 10,000 samples each, generated under null hypothesis of no lateral bias (Hit isoprobability across all positions in the display, with T targets in T equispaced positions); lines show the fit obtained with Equations 1, 4.*

---

[4] We mean, patients in whom the spatial processing stages that we are wishing to measure are intact.

Technically, in each simulation we obtained the conditional distribution of *MPH* given a fixed *T* and a fixed *H*. Thus for instance, with *T*=100 and *H*=20 we randomized the positions of 20 Hits across 100 targets, computed the *MPH*, and repeated the procedure 10,000 times. The result is the distribution of *MPH*s that a subject without spatial biases who detects *exactly* 20/100 targets would obtain by repeating the experiment many times.

The good news was that the shape of *MPH* distribution was satisfactorily close to Gaussian provided that there were at least 3 Hits and 3 Omissions – by 'satisfactorily' we mean that the absolute Kurtosis parameter was estimated to be less than .5 (see Table WM-3).

		Parameter									
		σ					Kurtosis				
	Targets	10	20	50	100	150	10	20	50	100	150
Hits	1	.319					**−1.222**				
	2	.213					**−.664**				
	3	.163	.166	.166	.166	.168	−.457	−.371	−.38	−.393	−.394
	4	.132	.14				−.398	−.322			
	5	.107		.127			−.337		−.281		
	6	.087	.108				−.373	−.234			
	7	.07					−.46				
	8		.086					−.266			
	10		.07	.084	.088			−.127	−.096	−.177	
	12		.057					−.24			
	14		.045					−.225			
	15			.065		.072			−.103		−.034
	16		.035					−.352			
	17		.029					−.476			
	20			.052	.058				.02	−.047	
	25			.042					−.048		
	30			.034	.045	.047			−.037	.04	−.059
	35			.028					−.072		
	40			.021	.036				−.192	−.049	
	45			.014		.036			−.249		.049
	47			.011					−.435		
	50				.029					−.015	
	60				.024	.029				−.004	−.032
	70				.019					−.006	
	75					.024					−.053
	80				.015					.001	
	90				.01	.019				−.142	.037
	97				.005					−.406	
	105					.016					−.008
	120					.012					−.073
	135					.008					−.044
	147					.003					−.466

Table WM-3 *Standard deviations (σ) and Kurtosis of sample distributions (simulated with N=10,000) of MPH as a function of number of Targets and number of Hits. Kurtosis values in **bold** (i.e. whose modulus was higher than .5) were assumed to indicate serious departures from gaussianity.*

The expected, bad news is that the σ_{MPH} changes dramatically as a function of *T* and, especially, as a function of Hit rate, *H/T* (see Fig. WM-2). The relationship between *T* and σ_{MPH} is a simple one, and is perfectly described by the Central Theorem of Limits (CTL): σ_{MPH} shrinks as a function of the square root of *T*. σ_{MPH} also shrinks as a

function of *H*: trivially, the mean position of, say, 5 Hits is much more variable than the mean position of 50 Hits. However, the equation of this reduction is not obvious, and does not look like a simple CTL. The reason is the following. We would have obtained a perfect CTL situation if the position of a Hit had been completely independent of the position of any other Hit, that is, if a Hit had been 'free' to occur at any position in the field of targets, irrespective of the position of other Hits. However, this is not the case in neglect tests: positions of targets are manipulated by the experimenter and are not the effect of random sampling. Most typically, the experimenter administers a homogeneous field of target positions, and if targets in a position have already been hit, that position will not be (or will be less frequently) available to be hit again: further hits will be forced, or more likely, to occur elsewhere.[5] This constraint on the position of a Hit is very mild with just a few Hits, and very severe with many Hits. More exactly, the constraint is totally absent when *H*=1: a single Hit is *completely* free to occur anywhere in the field of targets (indeed in this situation the distribution is a perfect Discrete Uniform, with analytically known parameters), and maximal when there is only one position left, i.e., when *H*=*T*−1: here a Hit will be forced to occur in the only residual position. Hence the σ curve closely follows the CTL for very small *H*, and drops towards zero as *H* increases – till it gets exactly at zero, of course, when all targets have been detected (*H*=*T*).

We searched for an equation that could predict the value of σ_{MPH} as a function of *T* and of Hit rate (*H*/*T*) – a complex guesswork as to what exact shape in any component would produce a satisfactory fit. At the end of the job we did obtain an equation which provided a very good fit (Fig. WM-2). Equation 1 was derived from the CTL applied to both *H* and *T* factors, and including the analytic formula for σ_{MPH} in the liminal case *H*=1, obtained from the Discrete Uniform Distribution, that is, $D = [(T^2−1)/12]^{1/2}/(T−1)$:

$$\sigma_{MPH} = \frac{CF[T(1-W)]\sqrt{\frac{(150^2-1)}{3H}}}{300(T-1)} \qquad \text{[Equation 1]}$$

W was the weighting factor used to account for the dependence between Hits positions, which increases with *H* (Equation 2); its best-fit scalar values were: $j = 1.155$, $k = −.56$, $m = .563$.

$$W = j\left(\frac{H}{T}\right)^3 + k\left(\frac{H}{T}\right)^2 + m\left(\frac{H}{T}\right) - \frac{j+150k+150^2m}{150^3} \qquad \text{[Equation 2]}$$

CF is a further correction term that takes into account possible repetitions of the same horizontal positions during the test – e.g., a same target position might be presented multiple times during the testing session, or, in a cancellation task there might be targets that are vertically aligned, thus having exactly the same horizontal position. An additional set of simulations could show that if a set of *T* targets is organized in *G* clusters of *T*/*G* repeated positions each, and again the positions are in an equispaced array, then σ_{MPH} is higher – up to 1.7 times higher – than when the *T* targets are not clustered. In the simple case with a single Hit, if all the *T* targets are clustered in two only groups – two only positions, $\sigma_{MPH} = .5$. We used this fact to obtain what we called the *Clustering Factor*, $CF = (.5/D−q−1)x^2 + qx + 1$; in it, *x* is a measure of the degree of clusterization, going from 0, no clusterization, to 1, maximal clusterization (2 only clusters): $x = 2(T−G)/[G(T−2)]$; *D* is defined above; the best-fit scalar value is $q = .477$; *G* is the number of clusters. *G* is to be estimated on grounds of the empirical target distribution in the following way: two targets are assigned to a same cluster if they are less than one hundredth of the display size apart (display size = distance between the leftmost and rightmost targets); if *T*>51 the limit is replaced by $1/[2(T−1)]$ of the display size.[6] Hence, the final formula for *CF* turned out to be:

---

[5] One should perhaps specify: 'elsewhere' in space and/or time. Indeed the violation of the statistical independence assumption regards the statistical data format and not the precise experimental setup. It does not matter whether one has one target per horizontal position, or multiple targets per horizontal position, presented simultaneously (i.e. vertically separated like in a cancellation task) or separately in time (like e.g. in Posner's paradigm). These scenarios are identical from the statistical point of view: if some or all of the targets in a given horizontal position have been hit, that position will be less available or not available at all for further hits, yielding a violation of the independence assumption. Note also that we do not consider or model perseveration (see e.g. Gandola et al., 2013) in this work.
[6] Without this rule (e.g.) 200 perfectly equispaced targets, which of course should be classified as 200 mono-target clusters, would all be classified as belonging to a single cluster!

$$CF = \left(\frac{T-1}{\sqrt{\frac{T^2-1}{3}}} - 1.477 \right) \left[\frac{2(T-G)}{G(T-2)} \right]^2 + .477 \left[\frac{2(T-G)}{G(T-2)} \right] + 1 \qquad \text{[Equation 3]}$$

In the procedure, we took advantage of the fact that the function governing σ_{MPH} can be fit only in the domain from $H=1$ to $H=T/2$, i.e. when the number of Hits is less than or equals half the targets. The extension of the function to the domain $H > T/2$ can be done just mathematically; it is sufficient to consider that the *MPH* distribution of, say, 20 Hits among 50 targets is identical to the distribution of the Mean Position of 20 Omissions (*MPO*) among 50 targets. Provided that both *MPH* and *MPO* are standardized in the (−.5, .5) space and C-adjusted (see Section 5.3 'Adjusting *MPH…*' below), there is only a scaling factor between them: *MPH = –MPO[(T–H)/H]*. Hence, for every $H > T/2$:

$$\sigma_{MPH(H)} = \sigma_{MPO(T-H)} \, [(T-H)/H] \qquad \text{[Equation 4]}$$

meaning that, to obtain the σ_{MPH} for a given number H of Hits, one has to compute $\sigma_{MPO(T-H)}$, the standard deviation for the $T-H$ Omissions (this can be done by simply applying Equations 1-2 above and replacing every H with $T-H$) and multiply the result by the scaling factor $[(T-H)/H]$.

For the liminal cases where the distribution is too far from the Gaussian, that is, when either Hits or Omissions are less than 3, the Worksheet proceeds as follows. When exactly 1 Hit or 1 Omission is produced, the *MPH* (or *MPO*) corresponds to the position of that single item; according to H_0 that Hit/Omission had equal probability to occur on any target of the display, hence the *p*-value is easily obtained by computing the percentile of the position of that single Hit/Omission within the distribution of target positions. As for the cases where exactly 2 Hits or Omissions are produced, we mapped standardized *MPH* values to percentiles by using the (non-Gaussian) data from the Monte-Carlo simulations: our Worksheet includes a table of the simulations' percentiles from which *p*-values are obtained. The table is currently (June 2018) only available for the case of no clusterization (*T=G*), so, if there is some significant clusterization (*G*>1.5) the Worksheet does not give any *p*-value (also look at the decision trees in Figs. WM-6 and WM-8).

Since we wished our Worksheet for automatic computation to analyze data sets with up to 256 targets, we checked whether the above Equations (which were obtained from simulations with maximum *T*=150) accurately predict the σ_{MPH} obtained from new simulations with *T*=256, and indeed they do.

4.1. Should one run a *z* or *t* statistical test?

The σ_{MPH} provided by our statistical model (and by our Worksheet) is to be used in an ordinary *z*-score formula: $z = (MPH-0)/\sigma_{MPH}$. The expected value is zero because normal subjects are assumed to have zero average *MPH* (this was shown to be true on the Diller task, see Section 5.6 'Empirical confirmation…'). For achieving slightly higher power, the C-adjusted should be preferred over the LCR-adjusted *MPH* (see Section 5.3 'Adjusting *MPH…*'). The classical Gaussian distribution is to be used, and not a *t* distribution, because σ_{MPH} is to be considered as known from our model, and does not need to be estimated empirically.

Mathematically speaking, a perfect performance (*H=T*) should correspond to an unknown *z*-score, as both *MPH* and SD are zero (*z*=0/0); however, to give a correct diagnosis of no neglect, the Worksheet gives a default outcome *z*=0, *p*=.5 in this case.

4.2. *H* variation within a subject

While H is unknown before the experiment, our statistical model treats it as a known parameter (each simulation studied the permutations of *exactly* H Hits in T targets). Thus our Equation gives a SD for *MPH* that does not take into account the effects by H differences that would occur if a same subject repeated the test many times. If one is interested in the *MPH* distribution of a bias-free subject repeating the test many times s/he should consider that our SD underestimates this variation. Nevertheless, correct neglect diagnosis needs exactly the *H*-variation-free SD that our model provides. We deliberately used the *H*-conditional distributions, exactly because they partial out the effects by H differences on diagnosis; while of course the main source of variation in H is inter-individual accuracy differences, also intra-individual random fluctuations are present; our model partials out both and restitutes an unbiased diagnostic outcome. We could confirm that this is the case by running an additional simulation in which

3,210 performances by a single subject were generated in each of 12 different combinations of T=10, 20, or 50, and p(H)=.1, .5, .7, or .9 (with isoprobability across positions). Therefore, this time each virtual subject was 'allowed' to vary for H, which followed the binomial distribution with parameter p(H). The resulting distributions of z-scores had SD very close to 1, precisely 1.0149 (averaged across the 12 simulations). This value is virtually identical to that predicted by the approximation error of our Equation, 1.0138 (Fig. WM-2), which was obtained from simulations with *fixed H* values. Average False Positive Rate was .0563, very close to the nominal 5%. So, the fact that simulations were generated with variable instead of fixed H values did not change the goodness of fit of our model in any way – in other words our model successfully partials out the effects of H differences, whatever their source.

In the following, we discuss the effects of violations of the statistical model's assumptions, and explain how the Worksheet for automatic computation treats them and reduces their negative impact.

5. VIOLATIONS OF MODEL ASSUMPTIONS AND REMEDIES (IF AVAILABLE)

The above Equation for SD holds when the assumptions of the model are met (see Sections below for a list of violations and of their possible effects), and, in general, for data sets with at least 10 targets. *P*-values can be obtained via Gaussian approximation [z=(C-Adjusted *MPH*)/SD] provided that there are at least 3 Hits and 3 Omissions [$2<H<(T-2)$]. However, the number of caveats is high and richly structured – see Section 6 'Algorithms and decision trees' and Figs. WM-5 and WM-6 for some insight into this complexity.

5.1. Violations of the assumption that the lower asymptote of the curve (*floor*) is zero

This apparently natural assumption is actually quite intricate in its meaning. To explain this point we will start from plausible cognitive models and explore what their mathematical implications are.

The 3-parameter logistic curve in which *floor*=0 (*f* for short) is logically implied by the idea that there are two main types of processing involved in neglect tasks. The first type of processing stages are those which, when damaged, produce neglect – a lateral bias in performance. These processing stages clearly encode the stimulus's spatial position, and will henceforth be called 'Spatial Processing Stages', SPS. Other processing stages are those that influence performance, but do not encode spatial position (e.g. face recognition, letter recognition, word recognition, object recognition, etc.). Lesion to them do not produce lateral biases, but only non-lateral deficits, that is, impairments that are exactly equal all across space. These stages will collectively be called "extra processing stages", EP.

We assumed the simplest, most intuitive, model for how the two kinds of deficit combine. The idea is that SPS produce some spatial parsing of the stimuli – if we assumed the attentional theory of neglect to be true, we would say that SPS 'move the focus of attention across objects'. When damaged, the probability of successful processing by SPS, p(SPS), is assumed to follow a logistic curve with free *slope*, *location* and *ceiling* parameters. The *slope* and *location* parameters express damage to SPS. As we saw, the *ceiling* parameter of this p(SPS) curve is assumed *not* to reflect SPS damage – if it did, we would be confusing the effects of damage to SPS with the effects of damage to EP [p(EP) is a flat function whose only parameter, which reflects the degree of damage, is *ceiling*]. Of course the *ceiling* of the p(SPS) curve can well be below 1: back to the attentional theory, the probability that attention falls on a given object is unlikely to be exactly 1, even in normal subjects. As for the *floor*, it would be weird to assume that the probability for attention to fall on an object can never be smaller than some above-zero limit. So, it looks reasonable to assume that p(SPS) has a zero lower limit – a zero lower asymptote.

Now, p(SPS) – modelled by a logistic curve with three parameters, and p(EP), the flat function governed by a single parameter, are not directly observable. What we observe directly is the curve ruling p(Hit). The question then is, how does p(Hit) depend on p(SPS) and p(EP)? The simplest idea is the *multiplication law*:

$$p(\text{Hit}) = p(\text{SPS})p(\text{EP}) \qquad\qquad \text{[Equation 5]}$$

For a Hit to occur, a target must be spatially parsed (SPS) *and* processed by the other critical stages (EP) at the same time. Take again the attentional theory as an example, and suppose that the task is to report a letter that can appear anywhere in a display: in order to have a Hit, a letter must *both* be reached by the focus of attention (SPS) and be successfully parsed by the reading system (EP). Thus those probabilities must be multiplied. Note that such a multiplication law holds if SPS and EP are *independent* processes – their success probabilities do not depend on each other.

The multiplication law implies that p(Hit), the product of a logistic 3-parameter curve and a flat function, is again a logistic 3-parameter curve. Also p(Hit) has a zero lower asymptote.

A hidden implication comes to the fore. If p(SPS)=0, p(Hit) *must* be zero. Rephrasing the statement, if spatial parsing does not succeed – *if the attentional focus does not reach an object*, a Hit is impossible. That is the logical consequence of the multiplication law.

This is a reasonable assumption for many tasks. For instance, in cancellation tasks or any other task where the subject is required to reach out to, or to gaze an object, if the object is not included in the attentional focus it will never be reached/gazed, so a Hit will indeed be impossible.

However in some other tasks, especially those with a symbolic response (like a verbal yes/no, or a button press) that is prompted by the examiner/apparatus, a Hit might well occur even if a target was *not* reached by the attentional focus. One such example is a task of detection of single visual stimuli, with each trial being signaled by a beep sound, and with the subjects being required to say 'yes, I saw the target' or 'no, I did not see anything'. Subjects might well guess a 'yes' response even though their attention did not reach the target – thus a Hit would be achieved without attention. The multiplication law has been violated here. Indeed, a non-multiplicative law holds:

$$p(\text{Hit}) = p(\text{SPS})p(\text{EP}) + f[1-p(\text{SPS})p(\text{EP})] \qquad \text{[Equation 6]}$$

where f is the probability to guess a 'yes' response when no target was perceived.

This model is characterized by a lower asymptote (*floor*) that is higher than zero, and equals f. Thus, different subjects would have different lower asymptotes. A direct demonstration that f is > 0 would be derived by looking at catch trials: if, as in many of the above experimental settings, there are trials in which no target-stimulus was actually given (only the warning 'beep' sound was delivered), a proportion f of these trials would receive a 'yes' response (so f = False Alarm rate in catch trials).

The problem can be solved, and we are working towards a solution. However for what concerns *MPH* in its current form (June 2018) this is a serious problem. Indeed the higher f, the more distorted (underestimated) the absolute value of *MPH*. Underestimation of *MPH* leads to reduction of the statistical power (sensitivity) of the diagnostic test for neglect.

Table WM-4 below shows by what factor absolute *MPH*s are reduced for f varying from .05 to .5.

f	True neglect severity (parameter v = true *MPH*)											
	.01	.02	.06	.1	.16	.217	.27	.33	.37	.41	.43	.44
.05	.95	.95	.94	.94	.93	.914	.89	.86	.81	.7	.5	.25
.1	.9	.9	.89	.88	.86	.834	.8	.75	.66	.52	.32	.14
.2	.8	.79	.78	.76	.73	.69	.64	.57	.47	.33	.17	.07
.3	.7	.69	.67	.65	.61	.565	.51	.43	.34	.22	.11	.04
.4	.6	.59	.57	.54	.5	.455	.4	.33	.25	.15	.07	.03
.5	.5	.49	.47	.44	.4	.358	.31	.25	.18	.11	.05	.02

Table WM-4 *Reduction factor of MPH as an effect of guessing behaviour probability f (= floor = lower asymptote of the logistic curve = False Alarm rate on catch trials). For example, .86 means that the observed (absolute) MPH is 86% of the true value.*

Therefore the users of our Worksheet (as of June 2018) are advised *not* to rely on *MPH* if both of the following conditions hold:

(i)   a Hit is well possible even *without* attention on a target (i.e. when spatial selection/processing was ineffective);

(ii)   one suspects that the patient 'guesses' some 'yes' responses even when s/he does not perceive the target – direct evidence in favour of this hypothesis is provided by the presence of False Alarms among catch trials.

Note that some cancellation tasks (e.g. the Diller & Weinberg, 1977, Letter Cancellation task) do have the equivalent of 'catch trials', that is, distractor letters; however the fact that a patient might cancel some distractors (technically, some False Alarms) does *not* invalidate the multiplicative law, because condition (i)

above does not apply: a Hit on a cancellation task is reasonably impossible without attention. The model in this case would still be multiplicative:

$$p(\text{Hit}) = p(\text{SPS})p(\text{EP}) + p(\text{SPS})[1-p(\text{EP})]f'$$ [Equation 7a]

Here f' is the probability of making a reaching movement 'guessing' that a stimulus that was successfully processed by attention [$p(\text{SPS})$] is a target, when letter identification failed (e.g. because of dyslexia or amblyopia), an event that occurs with probability [$1-p(\text{EP})$]. As you see, when $p(\text{SPS})=0$, also $p(\text{Hit})=0$: the multiplicative law still holds, and can be expressed by simply developing Equation 7a:

$$p(\text{Hit}) = p(\text{SPS}) \{p(\text{EP}) + [1-p(\text{EP})]f'\}$$ [Equation 7b]

While a full list of the tasks that are likely (or unlikely) to violate the multiplication law would be virtually impossible to prepare, some hints are given here. Recall that guessing behavior is the key factor for violating the law.

-We already classified cancellation tasks, and more generally, tasks involving 'analogical' responses – e.g. reaching out to the target, as very unlikely to violate the multiplication law: guessing is virtually impossible to succeed.

-If the task is visual search but the required response is symbolic – for example, a display containing many letters (the targets) is shown and subjects have to report all of them – guessing behavior is in general unlikely to occur; anyway such a behavior could easily be detected by including only a subset of the letters of the alphabet as targets: a subject who is guessing would produce letters that are not present in that subset. Alternatively, one can choose targets that are virtually impossible to generate by guessing, e.g. three-digit numbers, or number-letter pairs, etc.

-It the task requires a symbolic response to single targets that were presented in separate trials, the critical factor is whether or not the response is prompted. Tasks in which the response is *not* prompted – subjects cannot exactly predict when a target will be delivered – are likely to obey the multiplication law. E.g. in static perimetry the subject is administered with a sequence of stimuli with variable inter-stimulus interval, and without warning signals (also see e.g. De Renzi et al., 1989). This substantially reduces the probability of guessing behavior, because it is quite unlikely to give a response (a button press) within the time window of a stimulus without perceiving it. By contrast, we already saw that if a prompt signal is given on each trial (e.g. a 'beep' sound or an explicit question like 'have you seen a target?') the subject is forced to give a response; in this condition, guessing behavior is relatively likely to occur and to produce Hits (violation of the multiplication law).[7] However again, if the categories of the prompted response are *many* and not just two – for instance, if the target of each trial is an isolated letter and the subject has to report it rather than just say 'yes' or 'no', guessing behavior would be very unlikely to occur, because guessing the correct letter out of 26 or so is a hopeless strategy.

Note that the simple *presence* of guessing behavior is not sufficient to deduce a violation of the multiplication law: guessing must also be *effective*, that is, likely to have produced Hits. For instance, if targets are three-digit numbers presented on screen and a patient utters non-existent targets, this guessing behavior is very unlikely to succeed, so no sizeable violation of the law is implied.

Table WM-5 tries to classify tasks and suggests whether violations of the multiplication law are to be expected.

---

[7] If targets are isolated red/blue stimuli and the task is to name the colour after the 'beep', a violation of the law is clearly present, with the difference that we do not have Hits and Misses here, but Correct/Incorrect. The lower asymptote f is necessarily .5 (chance level), leading to a constant underestimation of *MPH*. Similar considerations hold for '2afc', two-alternative-forced-choice experiments (e.g. Azzopardi & Cowey, 1997).

Response type	N of targets per trial	Response is prompted?	Response categories	Examples	Violation of the law
'Analogical' (arm/hand reaching movement, or eye movement)	One/Many	No/Yes		Cancellation; picking up objects from a table; copying several separate and different drawings; gazing at target objects in a visual scene	No
Symbolic	Many	No	Many	Visual search of numbers; free recall of objects previously presented in different positions	Unlikely; however use targets that are difficult to guess
	One	No	Few	Detection of isolated stimuli presented with variable ISI (e.g. static perimetry)	Unlikely
			Many	Reporting of isolated letters presented with variable ISI	Unlikely
		Yes	Few	Prompted Y/N detection of isolated stimuli; prompted classification of stimuli as red/blue (2afc); recognition (Y/N) of objects previously presented in different positions	**Likely in YN designs – use catch trials to check for effective violation; violation is necessarily present in 2afc designs (f=.5)**
			Many	Prompted reporting of isolated letters	Unlikely – use catch trials anyway

Table WM-5 *Classification of task characteristics in terms of probability of violation of the multiplication law. YN = Yes/No designs in psychophysics; 2afc = two-alternative-forced-choice designs in psychophysics.*

5.2. Violations of target homogeneity assumptions

Violations of the target homogeneity assumptions (equispacing and ties homogeneity) can be classified according to the spatial frequency of the fluctuations in target density. Take equispacing as an example. Positions might not be equispaced, but the inter-position interval might vary *randomly* along the dimension; in this way the density of targets varies with high spatial frequency ('random-distribution' violation). Else, suppose that inter-position intervals systematically increase, or decrease, along the dimension: for example, positions might be denser on the left than on the right half of the display (an example of what we shall call 'eccentric-mean' violation), or be denser at display centre than at display ends, or show a pattern of alternate high- and low-density regions: here we have low-spatial-frequency violations of equispacing. By far the worst kind of a violation, potentially inducing massive biases in *MPH*, is the 'eccentric-mean' violation, occurring when targets' mean position is not halfway across the display; by contrast, distortions in *MPH* induced by random-distribution violation are typically minuscule and can safely be ignored.[8]

Eccentric-mean violations were studied in detail – see the following Sections (5.3 'Adjusting *MPH* for eccentric…' and 6 'Algorithms and decision trees').

It is important to note that if target distribution is *seriously* eccentric, problems will not be just statistical. Subjects' attention will most likely be biased towards the denser side, thus, ν will be non-zero as a baseline; these would make a diagnosis of neglect and a quantification of it (how far ν is from baseline) much more complex an enterprise than that formalized in this work.

5.3. Adjusting *MPH* for eccentric target distributions

By 'eccentric'[9] we mean target distributions whose mean position (Mean Position of Targets, *MPT*) is not exactly halfway between the leftmost and rightmost targets. In these cases an adjustment is necessary, otherwise one would misdiagnose a subject who detects all of the targets as showing neglect. In order to decide how to adjust the *MPH*, we had a look at the effects of massive non-linearity in the spacing of target positions on *MPH*. We used a

---

[8] Identical considerations hold for the other assumption, ties homogeneity.
[9] We did not use the more intuitive term 'asymmetric' because there are asymmetric distributions whose mean is the display centre (e.g. target positions: 0, 1, 1, 2, 4, 4).

logarithmic distribution of 101 positions across a virtual display. By using logistic functions, we implemented a whole range of neglect severities, from extreme left neglect to no neglect to extreme right neglect, also varying the *slope* of the logistic function.

Fig. WM-3 shows the results of this simulation. The distorted *MPH* obtained from the logarithmically-eccentric target distribution is shown on the *X* axis, while the correct *MPH* – the one that would have been obtained with perfectly equispaced targets – is shown on the *Y* axis. Data from a very shallow (low *slope*) and a very steep (high *slope*) logistic function are shown separately. Clearly, when there is no neglect, the *MPH* corresponds to the *MPT* – in this case *MPT*=+.29. When neglect is not zero, the dots show the behaviour of the distortion, which is somehow non-linear, and with different slopes on the two sides; anyway, at both ends, the distortion disappears, as an extreme right neglect (only the leftmost target is detected) yields −.5 and an extreme left neglect (only the rightmost target is detected) yields .5 anyway.

If one is wishing to eliminate the distortion, an adjustment is to be introduced that makes *X* values correspond to *Y* values in the plot, by means of a function interpolating the cloud of points. The problem is that the curvature of that function depends on the exact type of eccentricity in the target distribution; we explored the logarithmic distortion, but of course, there are virtually infinite ways for a distribution to be eccentric. So we were forced to simplify the procedure by using two separate straight lines (shown in solid in Fig. WM-3), those connecting the point (*MPT*, 0) to (−.5, −.5) and to (.5, .5).
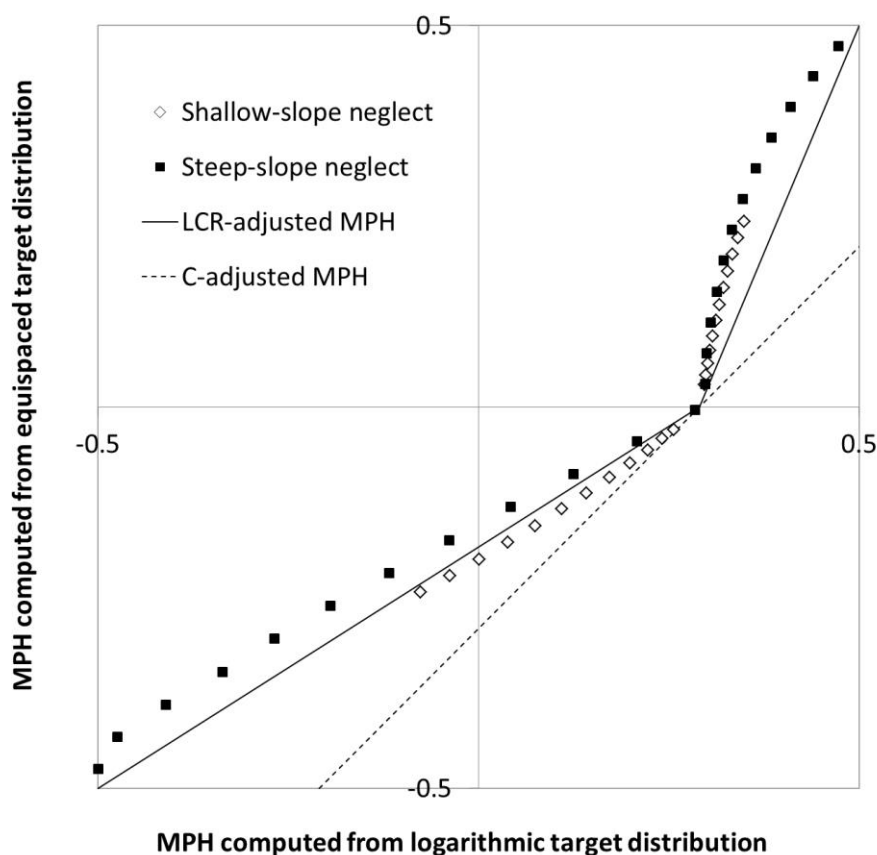
Figure WM-3 *Biases in MPH estimation produced by a skewed (logarithmic) distribution of target positions in space. Horizontal axis: MPH computed with the skewed distribution; vertical axis: correct MPH (the one that would have been obtained with equispaced target positions). 'Shallow'- and 'Steep-slope neglect' refer to the Hit-rate drop of the logistic curves describing neglect. Solid and dashed lines show, respectively, the LCR- and C- adjustments to MPH for reducing the biases.*

This adjustment corrects most of the discrepancy, albeit it is certainly suboptimal. We will call this adjustment 'LCR' (Left-Centre-Right), because it standardizes the *MPT*, which will always be zero, and both ends, with the leftmost target always being given the value −.5, and the rightmost target always being given the value .5.

There is a simpler way to adjust *MPH*, that is, to subtract the *MPT* position from the *MPH* position, and ignoring the ends. This will be called 'centering' (C-adjustment) and is shown as a dashed line in Fig. WM-3. C-adjustment was used, for instance, by Rorden & Karnath (2010) when computing their *CoC* index. What C-adjustment guarantees is that normal subjects (who detect all or almost all of the targets, hence having a *MPH* that is very close to *MPT*) will obtain a zero score; however, it does not consider the distortion induced when quantifying different degrees of neglect: the dashed line is indeed very far from the cloud of points representing the real performance (parameter ν) of neglect patients.

A simple simulation study confirmed that LCR-adjustment is better than C-adjustment, albeit still imperfect, when one is wishing to estimate the real level of neglect (ν) of a patient and target distribution is eccentric. The opposite holds true when one is wishing to *diagnose* neglect in the same situation: here the LCR-adjustment performs worse than the C-adjustment. Indeed, the distribution of LCR-adjusted *MPH* under null hypothesis of no neglect, is biased (its mean is different from zero), is often skewed, and has inflated standard deviation with respect to the

equispaced-targets case; the C-adjusted *MPH* has better distributions: these are unbiased (their mean is virtually always zero), are less often skewed, and are less often inflated in variance with respect to LCR-adjusted *MPH*.

To summarize, both C-adjusted and LCR-adjusted *MPH* scores are suboptimal corrections for target distribution eccentricity (Fig. WM-3). If one uses the C-adjusted *MPH* the bias will be eliminated when $\nu=0$ (no neglect), but will be larger, the more severe neglect (the farther ν is from 0). If one uses LCR-adjusted *MPH* there will be some bias also when $\nu=0$, but biases elsewhere will be smaller than with C-adjusted *MPH*. As a consequence, C-*MPH* is better for diagnosing neglect, and LCR-*MPH* is better for quantifying neglect. So in our Worksheet we used C-*MPH* in neglect diagnosis (*p*-value computation) and LCR-*MPH* for giving neglect severity estimates. This accounts for why sometimes the *z*-score does not exactly equal *MPH* divided by standard deviation: the *MPH* given as an output is the LCR-adjusted one, while in *z*-score computation the C-adjusted one is used.

While C- and LCR-adjusted *MPH* are 'specific' for diagnosis and quantification respectively, one must not forget that they are suboptimal: in cases of severe eccentricity of the target distribution, the *MPH* distributions will be irregular despite the use of adjustments. Hence the Worksheet has tools for diagnosing severe eccentricity in target distribution, which warn the user about the reliability of *MPH*. More in detail, when a severe eccentricity is caused by the spacing between targets being strongly asymmetric (see Section 6 'Algorithms…' below for details), the user is advised to give up inference on *MPH*, and to look at *MOH* instead. This index (Mean Ordinal position of Hits) only takes into account the ordinal positions of targets, and not their metric positions in physical space. Thus for instance, if physical target positions are 0, 1, 2, 10, 20, these are ranked 1, 2, 3, 4, 5, and the *MOH* of a subject omitting the two leftmost targets will be rank 4. Clearly, inferences with *MOH* are limited to the abstract space of target order.[10] While this second-choice *MOH* index is completely immune to violations of the equispacing assumption (because it ignores the metrics), it is not immune to serious eccentricity (and in general, to serious violations of homogeneity) of ties distribution. In other words, irrespective of whether positions are equispaced, when (e.g.) positions on the right have *many* more targets (repeats, or 'ties') than positions on the left, also *MOH* becomes unreliable.

It is important to note that the example in Fig. WM-3 was deliberately chosen for its being extreme, in order to illustrate the geometry of the problem. In the real world of neglect assessment, violations of that size are impossible, and in general sizeable violations of target distribution homogeneity are rare. In neglect experiments where targets are presented one at a time, their positions are typically *perfectly* equispaced and their number is *perfectly* balanced across positions. In visual search or cancellation tasks, the need for the display to have a disordered appearance (which in turn guarantees that target positions are to some degree unpredictable), forbids a perfectly equispaced organization of targets; however, target fields are almost always reasonably homogeneous and fall well within the limits that we set in our Worksheet for 'diagnosing' violations of the homogeneity assumption (and described in the 'Algorithms…' Section 6 below). Therefore the effects and discrepancies we described in the above macroscopic example are minuscule and negligible in practice.

5.4. Violations of the 'univocal interpretation' assumption

This assumption is very unspecific and shared by any possible measure of neglect. One example of violation is the behaviour of a patient without neglect who explores the display left to right during the test, and gets tired as time passes, thus missing progressively more targets. Here a right neglect would be misdiagnosed because of the confusion between time (the variable that really had an effect on detection probability) and space. Of course our model cannot avoid any such confusion, because it originates from an experimental rather than statistical limit of the procedure.

5.5. Violations of the isoprobability or no-lateral bias assumption

There are two possible violations of this assumption.
(1) Target detection probability varies *along the studied spatial dimension*: for instance (if the examiner is interested in the horizontal dimension) when targets on the left side are less likely to be detected than targets on the right side in a normal subject. In mathematical terms, this means that the logistic curve is not perfectly flat.

---

[10] If the user wishes to, s/he can directly insert ranks (ordinal values) instead of metric values (cm, mm, or any units of physical distance) as an input to the Worksheet; in this case, only *MOH* and no *MPH* solution will be given.

(2) Target detection probability within a given normal subject varies along *any* dimension that is orthogonal to (independent of) *the studied spatial dimension* – for instance, colour, shape, duration, time of exploration, … even *another spatial dimension*.

Type (1) violations tend to produce slight underestimation of the true *MPH* variability values, thus slightly increasing false positive rates (alpha probabilities in diagnostic decisions).

Type (2) violations produce the opposite effect, that is, a slight overestimation of true *MPH* variability with decrease in false positive rates. An example will clarify the latter case. Suppose that there are 50 red targets and 50 green targets, interspersed with 100 blue distractors, and that the subject is colour blind, so that he cannot tell green from blue stimuli. He will easily spot the red targets, but will miss all the green ones. According to our model, Hit rate=.5 predicts that *MPH* should vary, from sample to sample in the same subject, with standard deviation=.03; but we know that *MPH* will not vary at all: it will always be exactly 0 (no variation), because the red targets will always be detected, and the green ones will never be found. Hence, because of the violation of the assumption (red targets have higher detection probability than green ones), the *MPH* variability has been overestimated, and the test is over-conservative, with a decrease in the false positive rate. Less extreme colour blindness (e.g. green targets are detected with 50% probability) will produce intermediate results, but always with some degree of over-conservativity. Another example can be fabricated by replacing 'red' with 'top' and 'green' with 'bottom', so that 'colour-blind' is replaced by 'vertical neglect'. Cases of vertical neglect without horizontal neglect would produce exactly the same over-conservativity of our model.

Some parts of the following Section were also included in the paper (Toraldo et al., 2017).

5.6. Empirical confirmation of the isoprobability assumption: normal subjects only vary for Hit rate, and have no lateral biases

The isoprobability or 'no lateral bias' assumption, which was introduced just to simplify the model's mathematics, has the following meaning: neurologically intact subjects and patients without neglect are supposed to vary only in the efficiency of those stages of processing that do not encode the spatial position of the stimulus (e.g. shape processing, if the targets need to be recognized by shape). This means that the sparse Omissions produced by neurologically intact subjects exclusively depend on occasional failure of space-invariant processing, with exactly equal probability in each and every position of the display. Of course, different subjects can have different levels of such spatially constant probability. In this view, all of the variation in *MPH* across normal subjects would actually be due to space-invariant factors.

We tested this assumption by looking at empirical data. The prediction is that the distribution of *z*-scores computed under the null hypothesis that no lateral bias is present, should distribute with mean=0 and standard deviation=1 in a sample of normal subjects (taking the *z*-score from each subject is a way of partialling out between-subjects differences in Hit rate).[11] We collected a sample of 199 controls (Female: 57%; age: 60.9±12.2, education: 11±4.6)[12] performing the Diller & Weinberg's (1977) letter cancellation task or variants of it – targets ranged 104 (as in the original) to 108, could be either Hs (as in the original) or Vs, and could be administered on A4 or A3 sheets. These data were found in the electronic archives of many different experimental or clinical studies in the lab of one of us (AT) across many years (1994-2013). Of the 199 subjects, 134 were excluded because they produced a perfect performance, detecting all targets, which leads to an unknown *z*-score (0/0); the distribution of the remaining 65 subjects' *z*-scores almost perfectly matched the standard Gaussian: the mean was −.024 (not significantly different from zero: one-sample *t*-test, $t(64)=.181$, $p=.857$, Bayes Factor *BF*=1083 against the hypothesis that normal subjects lie .5 standard deviations of 'pure noise' away from $z=0$) and the standard deviation was 1.077 ($\chi^2(64)=74.24$, one-tailed $p=.179$; Bayes Factor against hypothesis H_1 that variance was 2: *BF*=18.675).[13] Therefore, the assumption that no lateral biases affect a sample of 65 subjects (i.e., that all of the variation in their *MPH* is the effect of the expected noise due to non-perfect Hit rate) was confirmed. This is a direct confirmation of

---

[11] We did not test brain-damaged patients without neglect because the (very likely) inclusion of cases of subclinical neglect, no matter how small, would have polluted the evidence and made interpretation ambiguous; furthermore, what neglect test should have been used as an exclusion criterion? An inescapable circularity would have affected such an experiment.

[12] Demographics could be traced back for 76% of the overall sample.

[13] There was a clear outlier, with $z=3.318$; this subject missed 11 out of 108 targets, 10 on the left half and 1 on the right half of the display; the absolute deviation of his *MPH* was minor (+.03 or 3% of the display width); however, even excluding him on suspicion of some undetected minor brain damage (a legitimate move: H_1 specifies that the shape is Gaussian, albeit with σ>1, and not that there are outliers!), the group mean was −.076 ($t(63)=.611$, $p=.543$, *BF* = 9210) and the standard deviation was .999 ($\chi^2(63)=62.9$, one-tailed $p=.48$, *BF* = 225.12).

the validity of our statistical model, and of the neglect diagnoses yielded by it.

5.7. Gaussianity violations

When the number of positions is very small, there is an impact on the shape of the statistical distribution of *MPH*. Thus if 50 targets are presented in 50 different positions, the *MPH* distribution will be very close to Gaussian, while if the 50 targets are presented in only 2 positions, *MPH* may have a very discrete shape, also depending on the overall number of Hits/Omissions. Of course such situations are very rare in practice; however, when the shape is too irregular to consider the Gaussian approximation as reliable, a warning message is given by the Worksheet and the *p*-value is omitted from the Output.

As explained in previous Sections, with fewer than 3 Hits or 3 Omissions, *MPH* distributions importantly depart from gaussianity. However, when there is exactly 1 Hit, the distribution of *MPH*, that corresponds to the position of that single Hit, is nothing else than the distribution of target positions; hence the *p*-value is simply computed as the cumulative of all target positions that are more extreme than the recorded one. In this situation the *p*-value is valid irrespective of any form of model violation (e.g. violations of the equispacing assumption and/or of the ties-homogeneity assumption). Identical considerations hold for the case with exactly 1 Omission. When there are 2 Hits, or 2 Omissions, we obtained stable estimates of many percentiles of the (non-Gaussian) *MPH* distribution (see Table WM-4) – hence the Worksheet provides interpolated *p*-values; however these hold only when there are no ties – i.e. when each positions has only one target in it – warning messages are given otherwise.

1-tailed *p*	Two Hits						Two Omissions					
	T=10	T=20	T=50	T=100	T=150	T=256	T=10	T=20	T=50	T=100	T=150	T=256
.500	.000	.000	.000	.000	.000	.004	.000	.000	.000	.000	.000	.000
.490	.000	.000	.005	.005	.005	.006	.000	.000	.000	.000	.000	.000
.480	.000	.000	.010	.010	.010	.011	.000	.000	.000	.000	.000	.000
.470	.000	.026	.015	.015	.015	.016	.000	.001	.001	.000	.000	.000
.460	.000	.026	.020	.020	.022	.021	.000	.003	.001	.000	.000	.000
.450	.000	.026	.026	.028	.029	.026	.000	.003	.001	.001	.000	.000
.440	.056	.026	.031	.030	.034	.031	.014	.003	.001	.001	.000	.000
.430	.056	.032	.036	.038	.039	.036	.014	.004	.002	.001	.000	.000
.420	.056	.053	.041	.043	.044	.042	.014	.006	.002	.001	.001	.000
.410	.056	.053	.046	.048	.050	.047	.014	.006	.002	.001	.001	.000
.400	.056	.053	.051	.053	.055	.053	.014	.006	.002	.001	.001	.000
.390	.056	.053	.056	.061	.060	.059	.014	.006	.003	.001	.001	.000
.380	.056	.066	.066	.066	.067	.065	.014	.007	.003	.001	.001	.001
.370	.056	.079	.071	.071	.072	.072	.014	.009	.003	.001	.001	.001
.360	.056	.079	.077	.076	.079	.077	.014	.009	.003	.002	.001	.001
.350	.111	.079	.082	.081	.084	.084	.028	.009	.003	.002	.001	.001
.340	.111	.092	.087	.091	.091	.089	.028	.009	.004	.002	.001	.001
.330	.111	.105	.097	.096	.095	.096	.028	.012	.004	.002	.001	.001
.320	.111	.105	.102	.101	.101	.102	.028	.012	.004	.002	.001	.001
.310	.111	.105	.107	.106	.107	.109	.028	.012	.005	.002	.001	.001
.300	.111	.109	.117	.114	.114	.115	.028	.012	.005	.002	.002	.001
.290	.111	.122	.122	.121	.121	.122	.028	.015	.005	.002	.002	.001
.280	.111	.132	.128	.126	.129	.128	.028	.015	.006	.003	.002	.001
.270	.139	.132	.138	.134	.136	.135	.028	.015	.006	.003	.002	.001
.260	.167	.145	.143	.141	.143	.143	.042	.015	.006	.003	.002	.001
.250	.167	.158	.148	.146	.149	.150	.042	.018	.006	.003	.002	.001
.240	.167	.158	.158	.154	.158	.158	.042	.018	.006	.003	.002	.001
.230	.167	.158	.168	.162	.164	.164	.042	.018	.007	.003	.002	.001
.220	.167	.171	.168	.167	.169	.172	.042	.020	.007	.003	.002	.001

.210	.167	.184	.179	.177	.176	.179	.042	.020	.008	.004	.002	.001
.200	.222	.184	.189	.184	.186	.187	.042	.020	.008	.004	.002	.001
.190	.222	.197	.199	.192	.195	.195	.056	.022	.008	.004	.003	.002
.180	.222	.211	.199	.202	.201	.204	.056	.023	.009	.004	.003	.002
.170	.222	.211	.209	.210	.211	.212	.056	.023	.009	.004	.003	.002
.160	.222	.224	.219	.220	.218	.222	.056	.025	.009	.004	.003	.002
.150	.222	.237	.230	.227	.227	.229	.056	.026	.010	.005	.003	.002
.140	.222	.237	.240	.237	.235	.238	.056	.026	.010	.005	.003	.002
.130	.278	.263	.250	.247	.247	.248	.069	.028	.010	.005	.003	.002
.120	.278	.263	.260	.255	.255	.258	.069	.029	.011	.005	.003	.002
.110	.278	.265	.270	.263	.267	.268	.069	.029	.011	.005	.004	.002
.100	.278	.289	.281	.275	.279	.280	.069	.032	.012	.006	.004	.002
.090	.333	.303	.296	.285	.289	.291	.069	.032	.012	.006	.004	.002
.080	.333	.316	.311	.301	.300	.304	.083	.035	.013	.006	.004	.002
.070	.333	.316	.321	.313	.312	.320	.083	.035	.013	.006	.004	.002
.060	.333	.342	.332	.328	.326	.332	.083	.038	.014	.007	.004	.003
.050	.333	.355	.352	.343	.342	.346	.083	.038	.014	.007	.005	.003
.040	.389	.368	.362	.356	.361	.364	.097	.041	.015	.007	.005	.003
.030	.389	.395	.383	.374	.379	.381	.097	.044	.016	.008	.005	.003
.025	.389	.395	.398	.386	.393	.390	.097	.044	.016	.008	.005	.003
.020	.444	.421	.408	.399	.401	.403	.111	.045	.017	.008	.005	.003
.010	.444	.434	.434	.432	.431	.427	.111	.050	.018	.009	.006	.003
.009	.444	.447	.444	.432	.435	.429	.111	.050	.018	.009	.006	.003
.008	.444	.447	.444	.437	.438	.433	.111	.050	.018	.009	.006	.003
.007	.444	.447	.444	.442	.441	.439	.111	.050	.018	.009	.006	.003
.006	.444	.447	.449	.447	.445	.442	.111	.050	.019	.009	.006	.004
.005	.444	.461	.454	.452	.448	.449	.111	.051	.019	.009	.006	.004
.004	.444	.474	.459	.460	.451	.455	.111	.051	.019	.009	.006	.004
.003	.444	.474	.464	.465	.461	.463	.111	.053	.020	.009	.006	.004
.002	.444	.474	.469	.470	.468	.471	.111	.053	.020	.010	.006	.004
.001	.444	.474	.485	.482	.480	.479	.111	.053	.020	.010	.006	.004

Table WM-6 *Absolute values of MPH (use C-adjusted ones from the experiment) corresponding to one-tailed p-values (leftmost column), to be applied when exactly 2 Hits or 2 Omissions are recorded. Linear interpolations are computed by the Worksheet for intermediate number of targets (T).*

6. ALGORITHMS AND DECISION TREES

In this Section we report the algorithms and decision trees implemented in our Worksheet. By 'algorithms' (Fig. WM-4) we refer to the sequences of operations applied on the targets' coordinates to obtain *MPH* (or *MOH*) and SD estimates. By 'decision trees' we refer to one specific sector of such algorithms, that is, to the complex criteria for giving, or avoiding to give, SD and *p*-value estimates on grounds of possible irregularities in the target distributions (violations of the target homogeneity assumptions, Section 5.2).

6.1. The algorithms: 'metric' vs 'ordinal' target coordinates and 'metric' vs 'ordinal' solutions

Fig. WM-4 shows the main streams of information processing implemented in our Worksheet. The input is the set of coordinates in one dimension (we will call it 'vector *X*', albeit that might well be in the vertical dimension of course). The aims of the algorithms is to produce estimates of *MPH* (or *MOH*) and SD, from which *z*-scores and *p*-values are then generally computed.

METRIC COORDINATES

ExA: (**0,2**,499,501,998,1000), real MPH=-0.499   ExA: (**1,1**,500,500,999,999)   ExA: (**1,1**,2,2,3,3)
ExB: (**0,2**,49,51,998,1000), real MOH=-0.495   ExB: (**1,1**,50,50,999,999)   ExB: (**1,1**,2,2,3,3)

```
Original coordinates X  ──→  Clusterization ──→ Clusterized coordinates X' ──→ Ranking ──→ Clusterized ranks X''
```

cluster count G   cluster count G

```
        ┌──────────────────────────┐        ┌──────────────────────┐        ┌──────────────────────────┐
        │  MPH           SD        │        │ r'overall, r'spacing, r'TIES │ invalidate │  MOH           SD        │
        │      Metric solution     │        │ MPT, MPP, TIES │ or validate │      Ordinal solution     │
        └──────────────────────────┘        └──────────────────────┘        └──────────────────────────┘
```
invalidate or validate

estimated MPH=-0.499   **estimated MOH=-0.5**

ORDINAL COORDINATES

ExC: (**-3,-2**,-1,1,2,3), real MOH=-0.4   ExC: (**1,2**,3,4,5,6)

```
Original coordinates X  ──────────→  Ranking  ──────────────────────→  Ranks X''
```

cluster count G

```
        ┌──────────────────┐  invalidate   ┌──────────────────────────┐
        │  r' TIES,        │  or validate  │  MOH           SD        │
        │      TIES        │               │      Ordinal solution     │
        └──────────────────┘               └──────────────────────────┘
```

Fig. WM-4

estimated MOH=-0.4
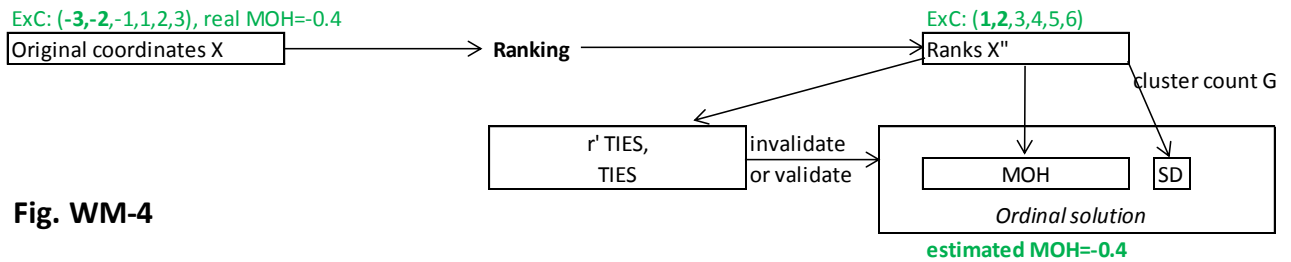
As the reader can see, the main distinction is between two cases: when coordinates are 'metric' and when they are 'ordinal'. By *'metric' coordinates* we refer to numbers that express real positions in physical space, and are thus given in some metric units like mm, cm, degrees, pixels, etc. By *'ordinal' coordinates* we refer to numbers that do not express measurement units, but just express a spatial order. For instance, suppose that in an experiment many stimuli were presented in each of three main sectors of space, Left, Centre, and Right, but the information as to the precise position within each sector was lost – only Left, or Centre, or Right are known for each trial. These horizontal coordinates are to be considered as e.g. –1, 0, +1, at an ordinal scale level.

When coordinates are metric, both a *metric solution* (*MPH*) and an *ordinal solution* (*MOH*) are offered; when coordinates are just ordinal, only the ordinal solution is meaningful, so it is the only one that is offered. Beware not to confuse the meaning of coordinates (which can be either metric or ordinal) with the type of solution (which can also be either metric or ordinal).

Clearly, if coordinates are metric, for instance expressed in mm, the *MPH*, that is, the *metric solution*, is fully viable – the metric meaning of the coordinates allow for averaging out the positions of the Hits and get a meaningful summary statistic, the *MPH*. Equally obvious is the case where coordinates are ordinal: in this case averaging out the coordinates (*MPH*) potentially leads to meaningless numbers – so the metric solution is prevented. What we do in this case is to take the ordinal coordinates and transform them into ranks, which can be averaged out (*MOH*) – we follow the *ordinal solution*. This new quantity can be interpreted just in the ordinal scale, i.e., in the abstract space of target order, without reference to real physical metrics.

So far, so good. The last, intriguing case is when coordinates are metric, and one wishes to look at the ordinal solution. This might seem odd – why should one wish to lose usable information, i.e. the exact spacing between positions? However, there is one particular condition in which this is fully justified, and even necessary. If targets are distributed in an irregular way (see 5.2 'Violations of target homogeneity assumptions') – so irregular that the statistical models, which were simulated by assuming equispaced positions and equal number of targets per position, cannot be applied, the ordinal solution is the only one that can be interpreted. We shall discuss the diagnostics of target distribution irregularities, which will sometimes explicitly deny the validity of the metric solution, and, in some extreme cases, even deny the validity of the ordinal solution.

It must be clarified here that in practice, ordinal coordinates are likely to be rare; even rarer is the possibility that target are distributed so irregularly as to make our statistical model invalid. Hence, the metric coordinates – metric solution combination will cover the vast majority of practical cases, However, for the sake of completeness, we are discussing all other cases here.

In the Worksheet (from version 2.0 onward) the default is that coordinates have metric meaning, unless the user overtly specifies that this is not the case – by plugging an 'o', for 'ordinal' in a dedicated cell.

Therefore, if the user specifies that coordinates are metric (or does not specify anything), both a *metric solution* (*MPH* plus statistical model) and an *ordinal solution* (*MOH* plus statistical model) are offered (top panel in Fig. WM-4); else, if the user explicitly specifies that the coordinates are just ordinal, only the *ordinal* solution (*MOH*) will be made available by the Worksheet (bottom panel in Fig. WM-4).

In more detail, the information processing is as follows. Fig. WM-4, top panel, shows the case were coordinates are metric. The original metric coordinates, *X*, give rise to a direct estimate of *MPH*; coordinates are then ***clusterized*** – i.e. targets whose coordinates are very close to each other (in general, less than 1/100 of the display width apart, see Section 4) are assigned to a same 'cluster', and are all given the average coordinate of that cluster. E.g., if targets whose coordinates are 1, 2 and 3 mm compose a cluster, new 'clusterized' coordinates are 2 mm for all three targets. The count *G* of clusters (see Section 4, text relative to Equation 3, for details on the computation) is used to derive the SD of *MPH*. Clusterized coordinates *X'* are then used to assess the possible presence of important violations of the target homogeneity assumptions. These assessment is made on ground of six parameters, called $r'_{overall}$, $r'_{spacing}$, r'_{TIES}, *MPT*, *MPP*, and *TIES*, which are all extensively explained in the Sections 6.2.1 and 6.2.2. The result of this assessment can be (i) invalidation of the metric solution, with validity preserved for the ordinal solution, or (ii) invalidation of both solutions, as will be explained in Section 6.2.3. As a last step, clusterized coordinates *X'* are transformed into ranks, i.e. the metric value is lost and a pure order, 1, 2, 3… is generated. On grounds of these ***clusterized ranks X"***, both *MOH* and its SD, composing the ordinal solution, are determined.

The bottom panel of Fig. WM-4 shows the information stream that is applied when original coordinates *X* are stated to be ordinal in nature by the user. In this case coordinates are immediately transformed into ranks, which give rise to MOH and SD – an ordinal solution, as before. The ranks are also used to derive two parameters, r'_{TIES} and *TIES*, which are useful to assess the reliability of the ordinal solution (the only one available in this ordinal-coordinates case).

6.1.1. *Comparison between ordinal solutions when coordinates are metric vs ordinal*

An *ordinal solution* is given both when coordinates are metric and when coordinates are ordinal (Fig WM-4). However, the mathematics is slightly different in the two scenarios. The differences lie in the fact that the metric nature of coordinates allows for some operations that the ordinal nature cannot justify. Indeed, since ordinal coordinates do not express true distances, one of the criteria for assessing target distribution homogeneity, i.e. equispacing (parameters $r'_{spacing}$, *MPP*) will not be meaningfully evaluated. Moreover, in the absence of distances, out of the six parameters that evaluate target distribution homogeneity, only *TIES* and r'_{ties} carry meaningful information. Target distribution homogeneity will hence be judged only with reference to these two scores. Also, the procedure used to count the clusters is different: in the case of metric coordinates, targets that were very close to each other (e.g. 1 mm apart on an A4 display) can be stated to belong to a same cluster; clearly, no such reasoning can be applied to ordinal coordinates: in this case any difference, no matter how small, will be evidence that two targets belong to different clusters. Thus, the number *G* of clusters will equal the number of different coordinates in the set. As a consequence, *G*, and hence *T/G* (the average number of targets per cluster, an important parameter, see later), will in general be different between metric and ordinal coordinates; also *TIES*, r'_{ties}, SD(*MOH*), *z*-score, and *p*-values, which all depend on *G*, will be slightly different.

6.1.2. *Clustering and computation of MPH/MOH and SD in various cases*

In this Section we will show some examples of computation carried out on coordinates, in order to illustrate the wheels and gears of the Fig. WM-4 schema, and to explain why it was built so.

Example A in Fig. WM-4. Consider the case of ***metric*** coordinates *X*=(**0,2**,499,501,998,1000) in which only the first two targets, *X*=1, *X*=2 have been detected (**bold**). An unbiased estimate of *MPH*, taken from the original coordinates, is (mean(0,2)−0)/(1000−0)−.5=(1/1000−.5)= −.499. To compute SD, which can vary by a factor of up to 1.7 because of clusterization (Section 4, text referring to *CF*), coordinates need to clustered first, and become *X'*=(**1,1**,500,500,999,999).[14]

---

[14] The small distortion in SD due to the discrepancy between original *X* and clusterized *X'* is many times smaller than the distortion we would have had by assuming that the positions in *X* were equispaced [*X*=(0,200,400,600,800,1000)]!

Example B. Now consider another example: *metric* coordinates are $X=(\mathbf{0,2},49,51,998,1000)$. This time the distribution is massively asymmetric, with severe failure of equispacing, and with a clear cluster-structure. We cannot use our statistical model (with Equations listed in Section 4) on this X, because that was simulated on equispaced positions (this invalidation occurs via the r' etc. parameters). So we need to transform X in order to make positions as close to equispacing as we can, and, at the same time, preserve the cluster-structure, otherwise SD will be severely distorted in its estimation. Trying to apply some form of curvilinear transformation is not a solution, because it would completely lack generality (it would need to be tailored on each specific case). The best, most practical, and most general way to proceed is to cluster coordinates first, and to rank them afterwards (the opposite operation would not work, see later). Thus we get $X'=(\mathbf{1,1},50,50,999,999)$ and from this, $X''=(\mathbf{1,1},2,2,3,3)$: a perfectly equispaced distribution, and which kept much of the cluster-structure. It is from this, and only from this X'' distribution that we can derive parameters *MPH* and SD, because X'' perfectly matches the feature (equispacing) of the distributions we used in our simulations to build the statistical model. So we have to derive *both MPH* and SD from $X''=(\mathbf{1,1},2,2,3,3)$. *MPH* of course is to be named *MOH* given that the new coordinates are ranks.

Example C in Fig. WM-4 shows a trivial case in which coordinates are ordinal.

A number of features can be grasped by inspecting the structure of the algorithm in Fig. WM-4.

First, the metric solution's *MPH* is computed from original coordinates, while SD is computed taking into account the number of clusters (i.e., conceptually, taking the clusterized coordinates). Indeed these two choices minimize estimation errors: clusterized coordinates would slightly bias *MPH* towards the extremes, and original coordinates (failing to take clusterization into account) would strongly bias the estimation of SD by a factor between 1 and about 1.7. As for the small bias in *MPH* estimation, this can be understood by comparing the estimation from the original coordinates X and the one obtained from the clustered X' coordinates. By using X', we would have obtained $(\text{mean}(1,1)-1)/(999-1)-.5=-.5$, which is slightly off correct: it indicates the left endpoint of the display, while we know that the detected targets were the leftmost, plus the leftmost-but-one (*real MPH* = $\nu = -.499$). By using X, we obtain the correct $-.499$ value. As for the bias in SD estimation if factor G, number of clusters, is not taken into account, see Section 4 (text referring to *CF*).

Second, the ordinal solution when coordinates are metric gives imperfect estimates, but no better estimates are currently available. Look at Example B in Fig. WM-4: at the end of the process we obtain *MOH* = $-.5$, i.e. the left extreme, while we know that the 'real' value is actually slightly off the left extreme, somewhere around $-.495$[15]: $X=(\mathbf{0,2},49,51,998,1000)$; it is anyway generally a very small bias and is well counter-acted by the possibility to apply the statistical model for diagnosis. Note also that, had we skipped the clustering operation, we would have got a much worse mistake: $X''=(\mathbf{1,2},3,4,5,6)$ would have led to *MOH* = $-.4$, much farther from the real (about) $-.495$ value than our current estimate, $-.5$.

Third, when coordinates are metric, target distribution homogeneity indices, r' etc. are computed from clusterized and not from original coordinates. Were they computed from the original coordinates, they would be distorted: the graph plotting metric against ordinal position, from which r' indices are conceptually derived, would look staircase-like when the coordinates of the targets of a cluster are separated by minuscule differences. For instance, try to plot the raw coordinates: $(1, 1.01, 2, 2.01, 3, 3.01, 4, 4.01, 5, 5.01)$ against their ranks $(1, 2, …, 10)$: the (staircase-like) plot is very far from the identity function, and this strong deviation is witnessed by r' parameters, even though the coordinates are really close to the perfectly regular $(1,1,2,2,3,3,4,4,5,5)$ distribution. This is why clustering is needed *before* computing r' in order for it to faithfully reflect the degree of regularity.

Fourth, when coordinates are metric, there is a difference between solutions. In the metric solution, we have seen, *MPH* and SD are computed from *different* coordinates (original and clusterized, respectively); in the ordinal solution, *both MPH* (which becomes *MOH*) and SD are computed from the *same* coordinates, the clusterized-and-

---

[15] There is no *precise* knowledge about were the 'true' *MPH*, i.e. ν, is in this case. Indeed, we know that *MPH* is an unbiased estimator of ν when target distribution is perfectly regular, and here the distribution is very far from regularity – it is massively skewed. However we do know that ν is not $-.5$, because the patient did not just detect the leftmost target, and it cannot be $-.4$, because that would have corresponded to a smooth, continuous violation of equispacing without *any* clustering at all [e.g. with an exponential scale $X=(0,12,43,129,364,1000)$]. We are certainly much closer to the clusterized case (we have three pairs of very close targets) than to the unclusterized, curvilinear-scale distribution, so a good guess is a value that is much closer to $-.5$ than to $-.4$, hence the $-.495$ value.

ranked ones. SD needs the clusterization operation otherwise it might be biased by the 1-1.7 factor mentioned above, and needs ranking otherwise the statistical models used in our simulations, and which assumed equispaced positions, would not hold. *MOH* needs the ranking for the same reason (otherwise the equispaced statistical model would not hold), and needs clustering otherwise a large centripetal bias (in Fig. WM-4's Example B, −.4 instead of −.495 or so) rather than a small centrifugal bias (in the same example, −.5 instead of −.495) would be chosen.

Fifth, when coordinates are metric, clusterization must be performed before ranking. If one reversed the order, ranking would cancel metric information altogether, so one would not be able anymore to understand which targets were metrically very close to each other, and could hence be put together in a same cluster. So, clusterization before and ranking afterwards is the only possible order. Look at Fig. WM-4's Example A: had we ranked the data before clustering them, we would have ranked X=(0,2,499,501,998,1000) into X'= (1,2,3,4,5,6), which would then have remained the same after clusterization, X''= (1,2,3,4,5,6), so the precious metric information of pairwise proximity between the targets would have been lost. Recall that this is a serious problem, because failure to acknowledge the clusterization structure in the data results in sizeable distortion in SD estimation. By contrast, if coordinates are *not* been metric, but jut *ordinal* (Example C in Fig. WM-4), no such proximity information exists in the data, so the direct ranking X'=(1,2,3,4,5,6) is completely justified.

6.2. Decision trees for 'metric' target coordinates

In this Section, we will explain down to fine detail how the criteria for assessing 'target distribution homogeneity' work when coordinates are metric (see the box containing the six parameters $r'_{overall}$, $r'_{spacing}$, r'_{TIES}, *MPT*, *MPP*, and *TIES* in the top panel of Fig. WM-4). Since sizeable departures from homogeneity of target distribution threaten the validity of the statistical model (and hence, of the diagnosis), such criteria are of crucial importance.
Decision trees regard the criteria used for deciding whether or not SD (Fig. WM-5) and *p*-value (Fig. WM-6) estimates are reliable, and for advising the user to rely on the metric solution (*MPH*) or rather on the ordinal solution (*MOH*).

Decisions are shown in the rightmost columns of the Figures, and are depicted in green when they are positive (SD or *p*-values can be estimated on grounds of our machinery) and in red when they are negative (our machinery does not have a solution). The terms reported in the Figures are clarified in the various Sections that follow.

Section 6.3 will show and discuss the decision trees (Figs. WM-7 and WM-8) to be used when target coordinates are ordinal.

The complex criterion for determining whether a target distribution is sufficiently homogeneous for guaranteeing reliability to our statistical model[16] was labelled '**general criterion of target distribution homogeneity**' in the decision trees (Figs. WM-5 and WM-6). To explain and justify such a criterion we need to summarize the kinds of violation to target distribution homogeneity.

6.2.1. r' indices

Departures from a perfect distribution of targets – the one that was used in the model's simulations – are easy to diagnose. It is sufficient to measure how far the distribution is from 'perfection', that is, an array of perfectly equispaced positions, each of which contains an identical number of targets (that is, the type of distribution we were able to simulate in our Monte Carlo study). One way to do so is to plot each metric position containing targets against its cumulative proportion of targets, and compute the Pearson correlation between them. The cumulative is the proportion of targets lying either to the left of the metric position, or *at* the metric position (half of the latter are considered in the computation). Thus, the (0,1)-standardized position X is plotted against $Y = proportion(t<X)+.5[proportion(t=X)]$, where t is the position of a given target) and compute the Pearson correlation r between X and Y. With the perfect distribution, all points lie on a straight line and r=1; any suboptimal distribution has r<1. We soon realized that with very small numbers of positions, for instance 2 or 3, r can be very misleading as it equals 1, or values very close to 1, even in cases of massively heterogeneous distributions (e.g., position A containing 10 targets, and position B containing a single target, yield r=1!). Hence we developed an alternative version of r, which we called r'. This new score does not consider the discrepancy of the points from the

---

[16] No Monte-Carlo simulations were run (yet) in this case, we just used intuitively prudent criteria.

Fig. WM-5

Fig. WM-6
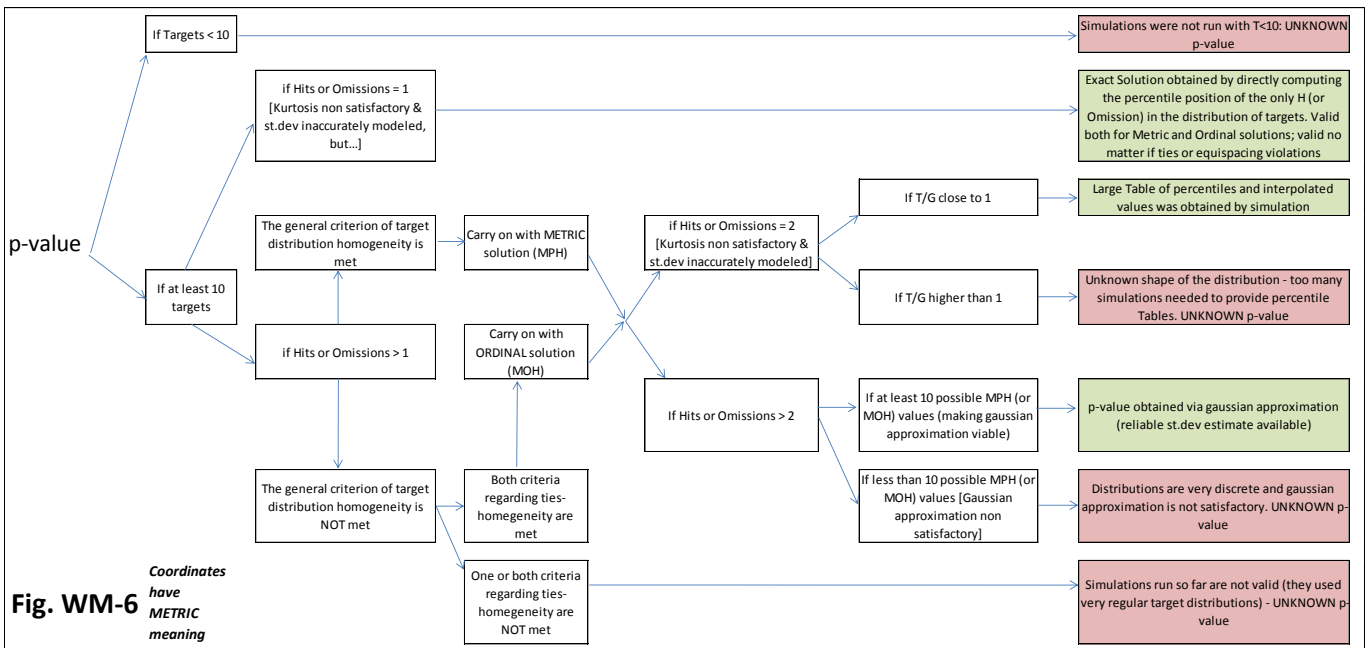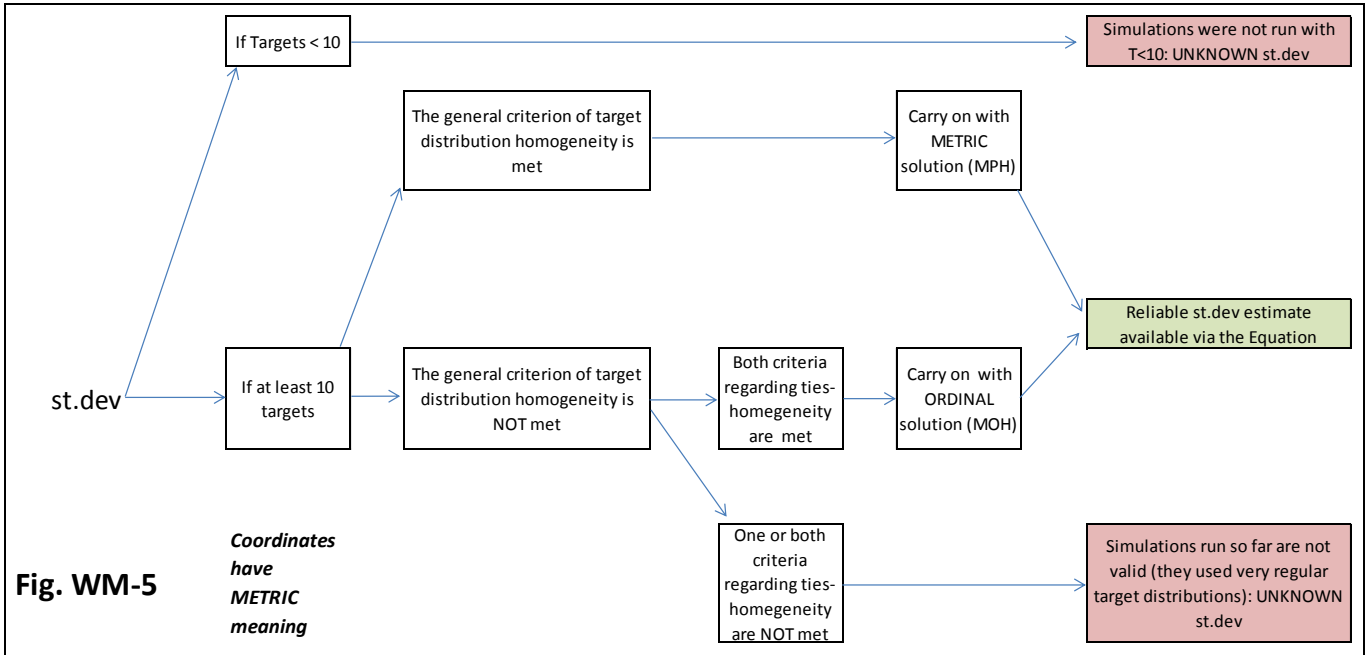
regression line, like *r* does; rather, *r'* considers the discrepancy of the points from the *identity function X=Y*. More precisely

$$r' = \frac{\sqrt{\frac{SS_{modelX}}{SS_{modelX}+SS_{noiseX}}}+\sqrt{\frac{SS_{modelY}}{SS_{modelY}+SS_{noiseY}}}}{2}$$

[Equation 8]

Where SS_{modelX} = Sum of Square deviations (in *X*) of the predicted points (those lying over the identity line) from their mean, SS_{noiseX} = Sum of Square deviations (in *X*) of the observed points from the identity line, and so on. Note that since the predicted points are over the identity line, $SS_{modelX} = SS_Y$, $SS_{modelY} = SS_X$ and $SS_{noiseX} = SS_{noiseY}$.

Also *r'* proved suboptimal when there are exactly 3 positions. Hence we derived a *r''* measure, to be used only when there are exactly 3 target positions. With *r''*, instead of taking the simple cumulative as *Y*, we take

$$cum'(X) = st.pos(X) + [obs.p.t(X) – exp.p.t(X)]$$

[Equation 9]

where *st.pos(X)*, or *standardized position of X*, is the metric position *X* rescaled to the (0,1) space (0=leftmost target position, 1=rightmost target position); *obs.p.t(X)*, or *observed proportion of targets in X*, is the proportion of targets in position *X* out of all targets; *exp.p.t(X)*, or *expected proportion of targets in X*, is the proportion of targets one

24

would expect if the distribution were perfectly homogeneous (i.e. *T/G*, where *G* = number of clusters = number of occupied positions).[17]

The *r'* measure is sensitive to any kind of departures from a perfect target distribution. We shall call it *r'*$r'_{overall}$, because it contains two separate components, which can be disentangled. What can be imbalanced and move *r'* away from 1 is either the distribution of positions themselves (equispacing violation), or the distribution of targets across positions (ties homogeneity violation). For instance, an imbalance in the distribution of *positions* is when targets are in positions (−5, −4, −3, −2, −1, 0, 2, 5): we have five positions in the left half of the display and only two in the right half (an instance of 'eccentricity'); or, (−100, −99, −98, −97, −96, 96, 97, 98, 99, 100): here positions are very dense at the far left and far right, and absent in the centre. Both cases are examples of violation of the equispacing assumption. Else, even though target positions are perfectly equispaced, the positions on one side might host more targets than the positions on the other side [e.g. targets in positions (−3, −2, −1, 0, 1, 1, 1, 2, 2, 2, 3, 3, 3), another instance of 'eccentricity']; or, lateral positions might host more targets than positions in the centre: (−3, −3, −3, −3, −2, −1, 0, 1, 2, 3, 3, 3, 3) – violations of the ties-homogeneity assumption, where by 'ties' we mean multiple targets presented at an identical position along the studied dimension.
Such a distinction is critical because the first type of violation, the one about equispacing, invalidates *MPH* and related statistics, but does not invalidate the non-parametric version of *MPH*, *MOH*. By contrast, violation of ties-homogeneity invalidates both *MPH* and *MOH*.
The disentanglement of the two components of $r'_{overall}$, which we named $r'_{spacing}$ and r'_{ties}, is not a difficult enterprise. For obtaining $r'_{spacing}$, we just erased the ties from the cumulative used in $r'_{overall}$ computation – thus the cumulative took into account each position as if it contained a single target. For obtaining r'_{ties} we just replaced the *X* metric positions, used in $r'_{overall}$ computation, with the *ordinal* positions [rank(*X*)].
Thus we obtained a triplet of indices, $r'_{overall}$, $r'_{spacing}$ and r'_{ties}, which are nicely interconnected and form a powerful and specific system of detection of target distribution irregularities (heterogeneities) of any kind. The three indices are the pairwise correlations of three variables: (i) *X*, the set of 'pure' metric positions (each position with multiple targets is counted only once), (ii) rank(*X*), the order of the *X* positions, and (iii) ***cum***, the cumulative proportion of targets across the *X* positions. Hence:

$$r'_{overall} = \text{corr}(X, \textbf{\textit{cum}}) \qquad \text{[Equation 10]}$$
$$r'_{spacing} = \text{corr}[X, \text{rank}(X)] \qquad \text{[Equation 11]}$$
$$r'_{ties} = \text{corr}[\text{rank}(X), \textbf{\textit{cum}}] \qquad \text{[Equation 12]}$$

This nice triangular correlational pattern is such that if one of the two components ($r'_{spacing}$ or r'_{ties}) is 1, $r'_{overall}$ equals the other component.
Before discussing the way to combine information from the three indices to obtain a single criterion, let us explore a second, parallel family of indices which capture one specific kind of violation in target distribution homogeneity.

6.2.2. *Indices measuring target distribution eccentricity*

Albeit the *r'* indices are exhaustive, in the sense that they capture any kind of departure from perfect homogeneity of distribution, there is one specific kind of departure that is likely to be particularly detrimental for our statistical model, so we introduced further indices to measure it. We refer to 'eccentricity' as we named it in the above Sections – a distribution is eccentric (or has an 'eccentric-mean' violation of homogeneity) when its mean is not exactly halfway between the positions of the two extreme targets. Also here we have three indices, which mirror the *r'* ones. The first, 'overall' index is obvious, and comes out directly from our definition of eccentricity: it is the location of *MPT*, the mean position of all targets, with respect to the geometrical midpoint of the display (i.e. the interval between the two extreme targets); by scaling the leftmost target's position as −.5 and that of the rightmost as +.5, the *MPT* should ideally be zero: if *MPT* is away from zero, this indicates *overall* eccentricity of target distribution (in analogy with $r'_{overall}$). Exactly as $r'_{overall}$, also *MPT* has two components, which correspond to the different possible causes of eccentricity in the distribution: one is when target *positions* are not centered at zero [e.g. (−5, −4, −3, −2, −1, 0, 2, 5): a violation of the equispacing assumption], another is when the frequency of targets across positions is not centered at zero, albeit positions *per se* might be perfectly equispaced [e.g. (−3, −2,

---

[17] For consistency with other analyses, before deriving all the indices discussed, *r'* , *r''*, *MPT*, *MPP*, *TIES* , we replaced the position of each target with the mean position of the cluster the target belongs to.

−1, 0, 1, 1, 1, 2, 2, 2, 3, 3, 3): a violation of the homogeneity of the distribution of ties]. Thus we have two more indices: the first is *MPP* [the Mean Position of (pure) Positions *X*], that is, the *MPT* after having excluded the ties (each position is counted only once irrespective of how many targets it contains), which quantifies the degree eccentricity due to the spacing of targets; by contrast, the degree of eccentricity due to ties distribution is assessed by looking at the *TIES* index, that is, the *MPT* that takes into account the ordinal positions rather than the metric positions of all targets. Both *MPP* and *TIES*, like *MPT*, range −.5 to .5. Thus we ended up with three new indices, *MPT* (the index for detecting and measuring 'overall eccentricity'), *MPP* (the index for detecting/measuring eccentricity due to spacing) and *TIES* (the index for detecting/measuring eccentricity due to the distribution of ties).

It is with these three indices that the logic of combination to produce the decision tree is easier to explain. A thorough analysis of the spectrum of possible eccentricity patterns in target distributions is reported in the following Sections 6.2.2.1 'Combining…' and 6.2.2.2 'Consequences…'. In short, if two of the three indices (including *MPT*) indicate sufficient centrality (i.e. are close enough to zero), interpretation of *MPH* in the classical, metric scale (reflecting physical space) and statistical tests regarding *MPH* are safe[18]; else, if the *TIES* index is close enough to zero, data can be interpreted non-parametrically, i.e. referring to the ordinal solution *MOH*, and all statistical tests regarding *MOH* are safe; if the *TIES* index is far from zero, we are getting too far from the realm of our simulations, and we are not sure what the behaviour (SD and *p*-value) of *MPH*/*MOH* is in those cases. Details as to the logic that led us to such a structure follow.

We arbitrarily set the limits for defining an index as 'close enough' to zero in this way: |*MPT*|<.05 (5% of the overall display width); |*MPP*|<.1; |*TIES*|<.1 (10% of the overall display width). Given that *MPT* is the most important indicator of eccentricity, this was required to be closer to zero than any of its *MPP* and *TIES* components.[19]

6.2.2.1. Combining different eccentric-mean violations

Overall eccentricity, as measured by *MPT*, expresses a combination of spacing eccentricity, as measured by *MPP*, and ties eccentricity, as measured by *TIES*. We shall refer to the latter two as the 'components' of overall eccentricity. If the two components are opposite (e.g. a leftward *MPP* and a rightward *TIES*), they tend to cancel each other out, and the target distribution is decently centered: our statistical model is still valid, *provided that the two components are not extreme*. Indeed the main factor invalidating our statistical model and Equations (see Section 4 'Monte Carlo simulations and model Equation') would be an *overall* eccentricity − hence if *MPT* is decently central [no more than 5% away from zero, .05 in the (−.5,.5) *MPT* scale], and the two components *MPP* and *TIES* are not too extreme (i.e., at least one of them is not farther than .1 from zero), this would validate our simulations, SD, *z* and *p*-values. If, by contrast, a close-to-zero *MPT* is the result of opposite, pathological components (e.g. *TIES*= −.3, *MPP*= +.3), the *MPH* sample distribution is unlikely to be regular and to be well approximated by our model which was based on perfectly regular target distributions.
This holds for the metric index of neglect (*MPH*). The ordinal index of neglect (*MOH*) is not influenced by spacing at all; hence *MPP* and its effects on *MPT* are irrelevant in this case: only the *TIES* parameter counts, and needs to be close-to-zero (not farther than .05 from it) for the *MOH* statistical model to be valid.
Therefore we decided that for the *MPH* statistical model to be valid, *MPT* (by the .05 limit) and at least one of the two components (*MPP* or *TIES*, by the .1 limit) must be close to zero. This guarantees overall centrality (*MPT*) with little variation between components (if one of them is close to zero, as *MPT*, variation between them must be relatively small).
By contrast, the validity of the statistical model for the ordinal neglect index (*MOH*) only depends on *TIES*, which needs to be close to zero (by a .05 limit).
Let us explore all the possible practical cases. To derive them, one needs to understand what the behavior of *MPT* is as a function of its two components. The relationship is not intuitive. If the two components are eccentric towards the same side (e.g. rightwards), *MPT* tends to reflect their sum [e.g. try positions (0,5,10,15,19,24,29,33,33,35), which have *MPP*=.039, *TIES*=.0375: *MPT*=.08]. If the two components are eccentric on opposite sides, *MPT* also tends to reflect their algebraic sum, thus falling somewhere in between them

---

[18] If *TIES* is the only criterion that is *not* met, beware that the ordinal solution (*MOH*) is not interpretable: only the metric (*MPH*) one is.
[19] Parallel limits (also arbitrary) were set for the *r'* indices: $r'_{overall}$>.975, $r'_{spacing}$ >.95, r'_{ties} >.95. When *TIES* and r'_{ties} are evaluated to decide whether the ordinal solution *MOH* is valid, they are screened with the stricter criteria, |*TIES*|<.05 and r'_{ties} >.975.

[e.g. (0,0,5,5,10,15,19,24,29,33,35), *TIES*= −.08, *MPP*= .04: *MPT*= −.04]. However if *TIES*=0 and *MPP*=.1, *MPT*=.1. Let us simplify with the following approximate mathematical rules:

(i) If one component (*TIES* or *MPP*) is zero, *MPT* equals the non-zero component (**un**like an algebraic sum);

(ii) if both components are eccentric towards the same side, *MPT* is more eccentric than both, on that same side (**like** an algebraic sum);

(iii) if the components are eccentric towards opposite sides, *MPT* falls somewhere in between them (**like** an algebraic sum).

Table WM-7 shows all possible combinations of the conditions of the three measures. The first classification is whether zero, or one, or two, or three of the measures are out of the 'close-to-zero' range [(−.05, .05) for *MPT*, (−.1, .1) for *MPP* and *TIES*]. **When 1 is out**, this may be either one component, or *MPT*. If the out-one is a component, the only possibility for this to happen is that it is *slightly* out (e.g. .12), the *MPT* is about .04, and the other component must be about −.08, on the opposite side (it cannot be 0, because if a component is 0 and the other is biased, *MPT* is automatically biased as the biased component, and we would be treating another case). When **2 are out**, the other one is implied to be 'in' – simplifying, 0. If the zero parameter is a component, we know that *MPT* automatically 'copies' the other, biased component (e.g. *TIES*=0, *MPP*=.12, *MPT*=.12). If the zero parameter is *MPT*, the only way to have the other ones out is to place them on opposite sides, like −.12 and .12. When **all 3 are out**, they are either all on the same side (in this case *MPT* is necessarily more extreme than both components) or two are on one side and the other on the opposite side (in this case, the two on the same side cannot be the two components, because if they were, *MPT* would be on their side too, and more extreme; so, the two components must be on different sides and the *MPT* must fall somewhere between them, but still out of the (−.05, .05) 'close-to-zero' range.

6.2.2.2. Consequences on the validity of the statistical model and Equations

Table WM-7 reports all of these possible topologies, and relative conclusions on the validity of the statistical model and Equations reported in the above Section 4 'Monte Carlo simulations and model Equation'. We specifically refer to the validity of the Equations used for computing SD, and to the validity of the Gaussian approximation assumed for computing *z*-scores (applied when there are at least 3 Hits or 3 Omissions).

Recall that validity of the metric-solution statistical model (*MPH*) requires *MPT* to be close to zero (with the .05 cutoff) and at least one component to be close to zero (with the .1 cutoff); validity of the ordinal-solution statistical model (*MOH*) only requires the *TIES* component to be close to zero (with the .05 cutoff).

Also recall that no systematic simulation study has been performed yet (the ways a distribution of targets can be distorted are virtually infinite, so this is a very demanding enterprise): the present indications are intuitive, and likely to be on the prudent side.

Metric-solution criteria	*TIES*	*MPT*	*MPP*	Validity of statistical Equations	
				Metric (*MPH*)	Ordinal (*MOH*)
0 out	0%	0%	0%	Yes	Yes
1 out (2 in)	−8%	5%	**10%**	Yes	**NO**
	12%	4%	−8%	Yes	**NO**
	4%	**8%**	4%	**NO**	Yes
2 out (1 in)	0%	**12%**	**12%**	**NO**	Yes
	12%	**12%**	0%	**NO**	**NO**
	−12%	0%	**12%**	**NO**	**NO**
3 out	**12%**	**24%**	**12%**	**NO**	**NO**
	−30%	**−18%**	**12%**	**NO**	**NO**

Table WM-7 *Validity of the statistical model and Equations as a function of all possible combinations of biases in the three indices of eccentricity of target distribution mean. Indices vary in the range (−50%,50%). Indices that are biased (away-from-zero) according to the criteria used for the validity of the Metric solution MPH (i.e. within the ±5% range for MPT, within the ±10% range for TIES AND MPP) are reported in* **red***; in one case, the TIES index is inside the ±10% range for validity of the Metric solution MPH, but it is outside the ±5% range for validity of the Ordinal solution MOH – this is reported in* **blue***.*

In the first column of Table WM-7, the number of 'in' and 'out' indices, according to the metric-solution criteria, are listed; nested under these broad cases, various combinations are shown.

When 1 index is off-zero ('out' in Table WM-7), we have three cases.

If only *MPP* is out, the ordinal-solution model is valid (equispacing is irrelevant here), unless the *TIES* index, which is 'in' according to the metric-solution criterion ($|TIES|<.1$), is 'out' according to the ordinal-solution criterion ($|TIES|<.05$). Metric-solution model is fine, because it depends more on the overall centrality (*MPT*) than on single components; the *MPP* bias is compensated for by the close-to-zero *TIES* that brings the *MPT* within close-to-zero limits.

If only *TIES* is out, again the metric-solution model is saved by a compensation of opposite effects: the eccentric *TIES* is counterbalanced by the opposite *MPP* eccentricity, leading to close-to-zero *MPT*. Two out of three are close-to-zero, including *MPT*, so the criterion for validity of the metric-solution model is achieved. By contrast, the ordinal-solution model is invalid: here only *TIES* counts, and this is biased (the effect of the spacing bias, *MPP*, which might have counterbalanced the *TIES* bias, is nullified because ranks, and not metric positions, are used with the ordinal solution).

If only *MPT* is out, both components must have small biases on the same side (otherwise the schema is impossible). The metric-solution model is invalidated by the eccentric *MPT*; the ordinal-solution model is valid because *TIES* is close-to-zero.

When 2 indices are off-zero, the metric-solution model is automatically invalidated, because it requires two indices to be close-to-zero. The case with close-to-zero *MPT*, and with *MPP* and *TIES* biased on opposite sides, is particularly interesting: this can be obtained for example, by extreme, funny distributions like: (0,0,0,0,0,0,0,0,0,0,100,101,102,103,104,105,106,107,108,109,110). *MPP* is .375, *TIES* is $-.214$, but *MPT*=0! The distribution is definitely weird, and does not guarantee that either statistical model would apply at all. As for the ordinal-solution model, this is valid when the only close-to-zero index is *TIES*.

When 3 indices are off-zero, neither model is valid.

6.2.3.   *Final criteria for signaling the invalidity of the statistical model*

On grounds of the study exposed in the previous Section, we prepared the following rules for establishing whether or not the statistical model is reliable.

   (i)   If two of the three eccentricity indices (*MPT*, *MPP*, *TIES*), including *MPT*, are close-to-zero (.05 range for *MPT*, .1 range for *MPP* and *TIES*), the metric-solution model and Equations (*MPH*) are valid.
   (ii)   If *TIES* is close-to-zero (.05 range), the ordinal-solution model and Equations (*MOH*) are valid.
   (iii)   Otherwise, *statistical* analyses (SD, *z*-scores, *p*-values) are unreliable: one can just look at absolute *MPH*/*MOH* values.

However, recall that three more general indices *r'* were proposed, which capture eccentricity as well as other forms of distributional irregularities. We reasoned that since eccentricity is likely to produce the most detrimental effects on the *MPH* statistical model, one way to give more weight to eccentricity in the decision about the model's reliability, is to combine eccentricity indices (*MPT*, *MPP*, *TIES*) with *r'* indices in the decisional structure (Figs. WM-5, WM-6):

   (i)   If two of the three eccentricity indices (*MPT*, *MPP*, *TIES*), including *MPT*, are close-to-zero (.05 range for *MPT*, .1 range for *MPP* and *TIES*), **AND** two of the three *r'* indices ($r'_{overall}$, $r'_{spacing}$, r'_{ties}), including $r'_{overall}$, are close-to-1 (.975 limit for *MPT*, .95 limit for *MPP* and *TIES*), the metric-solution model and Equations (*MPH*) are valid (this is the '**general criterion of target distribution homogeneity**' shown in the decision trees in Figs. WM-5 and WM-6).
   (ii)   If *TIES* is close-to-zero (.05 range) **AND** r'_{ties} is close-to-1 (.975 limit), the ordinal-solution model and Equations (*MOH*) are valid (this is referred to as '**Both criteria regarding ties homogeneity are met**' in the decision trees in Figs. WM-5 and WM-6).
   (iii)   Otherwise, *statistical* analyses (SD, *z*-scores, *p*-values) are unreliable: one can just look at absolute *MPH*/*MOH* values.

Note a feature regarding the criteria for validity of the metric-solution model (i). By analogy with the larger range we used for defining the components *MPP* and *TIES* as 'close-to –zero' ($-.1<MPP<.1$; $-.1<TIES<.1$) with respect to the range used for the overall *MPT* index ($-.05<MPT<.05$), we used a limit for defining the *r'* components as 'close-to-1' ($r'_{spacing}>.95$; $r'_{ties}>.95$) that was larger than the limit we used for the overall *r'* index ($r'_{overall}>.975$).

Recall that when there is one only Hit or one only Omission, *p*-values are obtained directly from the target distribution (Equations for deriving SD, and the Gaussian approximation, are not used), so all of the problems listed above regarding distributional irregularities do not apply.

6.2.4. *T/G and number of possible MPH/MOH positions*

In Fig. WM-6 *T/G* is the ratio between the number of targets *T* and the estimated number of clusters *G* (see Section 4: 'Monte Carlo simulations and model Equation' for details). Since *G* is estimated on grounds of the empirical target distribution, *T/G* may well not be an integer. In these cases *T/G* is rounded to the closest integer (this is necessary when, see Fig. WM-6, *T/G* is directly used, or when it is indirectly used to determine the number of possible *MPH* positions).
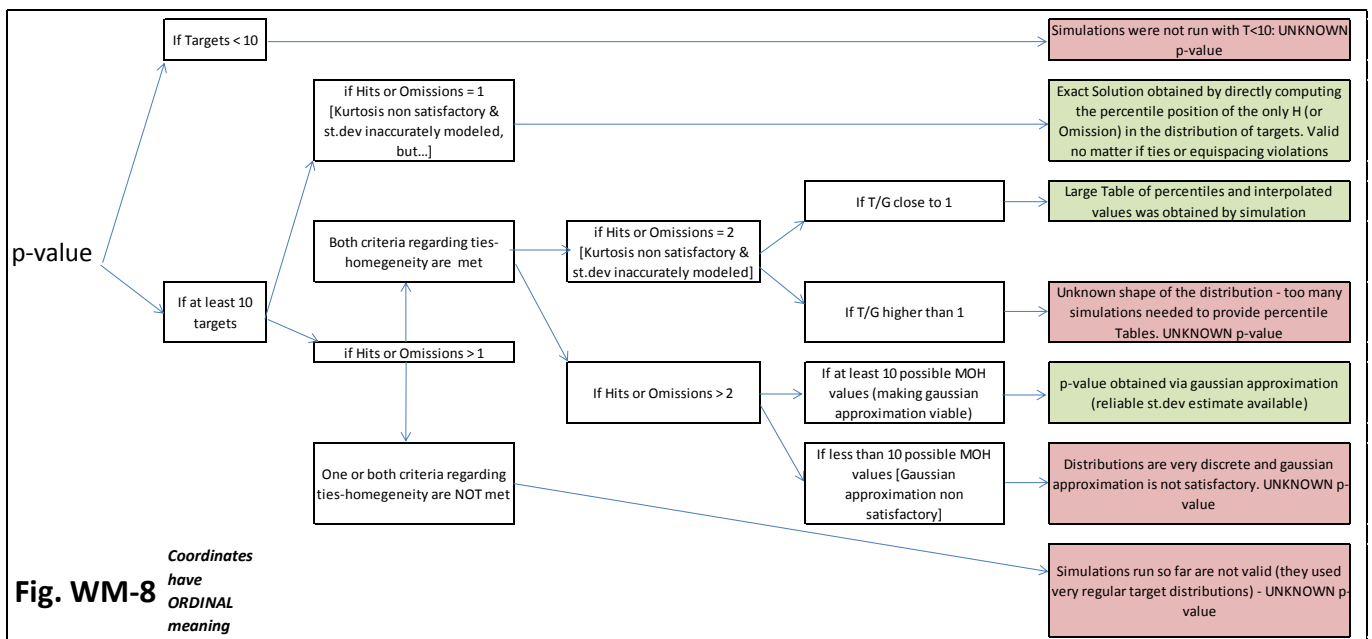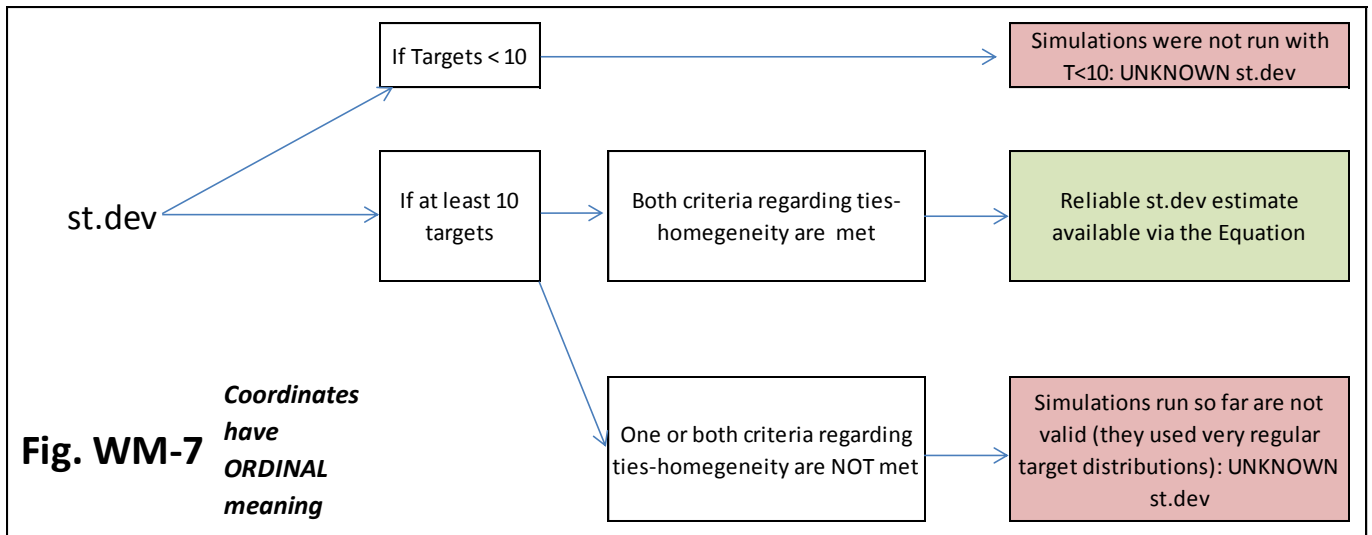
The number of possible *MPH* positions is a relevant parameter for deciding whether a Gaussian approximation is a good one. Suppose that *MPH* could assume just 3 positions (e.g. with 2 targets *MPH* can be −.5, 0, or .5): no continuous density function (like the Gaussian) could be meaningfully used to approximate so discrete a distribution. We arbitrarily set a limit at 10 discrete positions as the minimal number for accepting the Gaussian approximation. By running specific Monte Carlo simulations we could obtain the minimal Hit (or Omission) counts necessary for achieving the desired 10 discrete *MPH* positions, as a function of number of clusters *G* and of number of targets per cluster *T/G* (Table WM-8).

		Number of Targets per Cluster = T/G											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
N of clusters=G	**2**	-	-	-	-						10	11	12
	3	-	-	-		5	5	5	5	5	5	5	5
	4	-	-	3	3	3	3	3	3	3	3	3	3
	5	-	3	3	3	3	3	3	3	3	3	3	3
	6	-	2	2	2	2	2	2	2	2	2	2	2
	7	-	2	2	2	2	2	2	2	2	2	2	2
	8	-	2	2	2	2	2	2	2	2	2	2	2
	9	-	2	2	2	2	2	2	2	2	2	2	2
	10	1	1	1	1	1	1	1	1	1	1	1	1

Table WM-8 *Minimal Hit (or Omission) counts needed to have at least 10 MPH positions (a criterion for justifying the Gaussian approximation). E.g. if one has 5 clusters (row G=5) and 4 targets per cluster (Column T/G = 4), Hit count (or Omission count) must be at least 3 to satisfy the criteria. If G>10, the limit is always 1. If T/G>12, limits are identical to those for T/G=12. If T/G is not an integer, round it to the closest one. If a cell does not report a numerical limit, achieving 10 MPH positions is impossible. A '-' sign indicates that T<10, a condition that was never simulated (June 2018).*

6.3. Decision trees for 'ordinal' target coordinates

Figs. WM-7 and WM-8 below show the decision trees that are applied when coordinates have just an ordinal meaning. The terms are the same as explained in the previous Section 6.2 for the metric-coordinates case.
The general structure and logic is identical; the only difference is that there is no need to assess equispacing of target positions because the ordinal nature of the coordinates make this task meaningless. Hence two only of the parameters we introduced are now meaningful, that is, the ones assessing the distribution of ties: *TIES* and r'_{ties}. With these parameters, the same limits as those used for the most important scores in the metric-coordinates case (*MPT* and $r'_{overall}$), are to be applied: $−.05<TIES<.05$; $r'_{ties}>.975$).

Fig. WM-7 *Coordinates have ORDINAL meaning*

The figure shows a decision tree starting from **st.dev**:
- If Targets < 10 → Simulations were not run with T<10: UNKNOWN st.dev
- If at least 10 targets
  - Both criteria regarding ties-homogeneity are met → Reliable st.dev estimate available via the Equation
  - One or both criteria regarding ties-homogeneity are NOT met → Simulations run so far are not valid (they used very regular target distributions): UNKNOWN st.dev

Fig. WM-8 *Coordinates have ORDINAL meaning*

The figure shows a decision tree starting from **p-value**:
- If Targets < 10 → Simulations were not run with T<10: UNKNOWN p-value
- If at least 10 targets
  - Both criteria regarding ties-homogeneity are met
    - if Hits or Omissions = 1 [Kurtosis non satisfactory & st.dev inaccurately modeled, but...] → Exact Solution obtained by directly computing the percentile position of the only H (or Omission) in the distribution of targets. Valid no matter if ties or equispacing violations
    - if Hits or Omissions > 1
      - if Hits or Omissions = 2 [Kurtosis non satisfactory & st.dev inaccurately modeled]
        - If T/G close to 1 → Large Table of percentiles and interpolated values was obtained by simulation
        - If T/G higher than 1 → Unknown shape of the distribution - too many simulations needed to provide percentile Tables. UNKNOWN p-value
      - If Hits or Omissions > 2
        - If at least 10 possible MOH values (making gaussian approximation viable) → p-value obtained via gaussian approximation (reliable st.dev estimate available)
        - If less than 10 possible MOH values [Gaussian approximation non satisfactory] → Distributions are very discrete and gaussian approximation is not satisfactory. UNKNOWN p-value
  - One or both criteria regarding ties-homogeneity are NOT met → Simulations run so far are not valid (they used very regular target distributions) - UNKNOWN p-value

7. STATISTICAL POWER AND BAYES FACTORS

So far (June 2018) we provided the users of the Worksheet with diagnostic *p*-values. Computation of *p*-values has been relatively straightforward as *MPH* distributions proved to be nicely close to Gaussian. However these distributions concern *MPH*s obtained from perfectly flat Hit rate functions. In order to compute Power or Bayes Factors (*BF*) one would need to simulate the performance of several virtual patients with neglect, whose Hit rate function is very different from the constant function characterizing normal subjects. We expect such neglect-related *MPH* distributions to be skewed, thus a thorough, supplementary Monte Carlo study would be needed. However the main problem one has to face is that while the non-neglect 'curve' is known and univocal – it is the flat function varying only for the *ceiling* parameter (the height of the constant value), neglect curves are ambiguous, as they also vary for *slope* and *location*. Thus one *MPH* average value (one putative neglect severity level) corresponds to infinite combinations of *slope* and *location*, i.e. to very different *MPH* distributions.

We believe there is no theory-free way to specify what *MPH* distribution one should choose for Power/*BF* computation and why: a theory of neglect is necessary (see next Section 7.1 for examples). However our initial idea was that of keeping neglect theorization as far from our method as we could, in order to preserve its generality of application; thus we momentarily gave up developing a machinery for computing Power and *BF* because that would detract from the generality of the method.

7.1. Examples of effects by neglect-related theoretical constructs on parameters

There are several examples of different theoretical constructs one could make – here we only report some of them.

(i)   If a shift of the frame of reference were included in the pathogenesis of neglect (Karnath et al., 1998), this factor would directly influence the *location* parameter, which, however, would also reflect the extension of the attentional/representational scotoma. Moreover, if the degree of shift in the reference frame changed during the task, this variation would increase the final *slope* parameter.

(ii)   If the domain of the spatial representation that is damaged could shrink or enlarge, e.g. adapting to the size of the physical 'workspace' (e.g. the table, or the sheet, or the room, etc.), this variation would affect the *slope* parameter.

(iii)   If multiple spatial representations were used in a task, each of which has a very steep neglect gradient at different *location*s, the overall *slope* parameter would reflect the range of *location*s covered by the different representational components.

8. PASSWORD TO UNBLOCK THE CELLS IN THE WORKSHEET

If the reader wishes to look at the internal codes of the Worksheets (at least, of the versions that were published in 2017-2018), single sheets can be unblocked by using the following password: *sgrassatore*. The authors take no responsibility for the detrimental effects of changes in the code.

9. REFERENCES

Azzopardi, P., & Cowey, A. (1997). Is blindsight like normal, near-threshold vision? *Proceedings of the National Academy of Sciences*, *94*, 14190–14194.

De Renzi, E., Gentilini, M., & Barbieri, C. (1989). Auditory neglect. *Journal of Neurology, Neurosurgery, and Psychiatry, 52,* 613-617. doi: 10.1136/jnnp.52.5.613

Diller, L., &Weinberg, J. (1977). Hemi-inattention in rehabilitation: The evolution of a rational remediation program. In E. A. Weinstein & R. P. Friedland (Eds.), *Hemi-inattention and hemisphere specialization* (pp. 63–82). New York: Raven Press.

Gandola, M., Toraldo, A., Invernizzi, P., Corrado, L., Sberna, M., Santilli, I., …, Paulesu, E. (2013). How many forms of perseveration? Evidence from cancellation tasks in right hemisphere patients. *Neuropsychologia, 51 (14),* 2960–2975. doi: 10.1016/j.neuropsychologia.2013.10.023.

Gauthier, L., Dehaut, F.,& Joanette, Y. (1989). The Bells test: A quantitative and qualitative test for visual neglect. *International Journal of Clinical Neuropsychology, 11,* 49–53. Retrieved from http://strokengine.ca/assess/pdf/bellstest.pdf

Karnath, H.O., Niemeier, M., & Dichgans, J. (1998). Space exploration in neglect. *Brain*, *121*, 2357–2367.

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