UCL CENTRE FOR ADVANCED SPATIAL ANALYSIS

UCL

WORKING PAPERS SERIES

Paper 107 - Jul 06

Beyond statistical testing: Individual differences and the content and accuracy of mental representations of space

ISSN 1467-1298

casa

Centre for Advanced Spatial Analysis University College London 1 - 19 Torrington Place Gower St London WC1E 7HB
Tel: +44 (0)20 7679 1782 casa@ucl.ac.uk www.casa.ucl.ac.uk

Beyond statistical testing: Individual differences and the content and accuracy of mental representations of space

Victor Roger Schinazi

Centre for Advanced Spatial Analysis - University College London
Royal London Society for the Blind (RLSB)
v.schinazi@ucl.ac.uk

Abstract. The article uses data from two experiments on the content and accuracy of mental representations of space by the blind and visually impaired in order expose some of the shortcomings of typical statistical testing and propose an individual differences approach to the analysis of data. It begins with a discussion of some of the problems associated with the strict classification and eventual comparison of individuals between groups. The individual differences approach is then presented and the concepts of ability and present competence are explored along with the importance of detailed participant description. Examples from the two experiments are used to demonstrate how null hypothesis significance testing can be complemented with effect size estimates, box-plots and ranking techniques. Throughout the article we are reminded of the need to adopt mutually supportive techniques to account for the heterogeneity of experience and skills between participants.

Keywords: Mental representation, individual differences, visual impairment & blindness

1. Introduction

The disciplines of geography and psychology have a long tradition of comparing between different groups of participants. More often than not, research has focused on the way in which individuals within a group are similar to one another but different from individuals in another group. When comparing the performance of the blind and visually impaired in spatial tasks subjects are usually divided into different groups and a sighted control is assigned. Comparing the performance across groups can provide important information on the role of vision in the development of mental representations and can assist in the formulation of theories on sensory substitution and propioception. At times however, this inter-group method of comparison can be misleading especially when the adopted methodology does not allow for individuals to fully express their abilities. This is usually the case in tests that use blindfolded sighted controls. In many situations the control group is forced to operate at a disadvantage and rely on manipulatory and tactile strategies[1] that would not usually be used if vision were available. The same is true for blind and visually impaired subjects in experiments when the type, format and amount of information provided for the completion of a specific task favours the visual sense [1].

Not enough attention has been given to the ways in which individuals within a specific group differ from one another [2]. The phenomenological world of the blind and visually impaired is qualitatively different from that of the sighted [3]. Individuals with visual impairment and blindness form part of a population that is extremely heterogeneous that many times cannot be classified into specific groups or categories. The tradition of making comparisons between groups assumes that individuals that make up a particular group share the same characteristics. Individuals are grouped together because they have been diagnosed with the same eye or medical condition or have the same visual acuity. This type of classification can sometimes be restrictive. Consider the case of individuals who are diagnosed with the closest match condition. In such cases, the expert does not know the exact nature of the impairment and bases his diagnosis on the present manifestation of symptoms and behaviours. Some of the participants in this research have been diagnosed with a specific condition, most of the time retinitis pigmentosa

[1] Here we distinguish between primary and secondary strategies. Primary strategies are the natural, predisposed and frequently relied strategies used when problem solving. Secondary strategies are the best-adapted alternative when faced with a situation where the natural response cannot be elicited.

(RP)[2], but do not exhibit many of the characteristics that the condition incurs. This type of unforced professional error cannot account for latent behaviours or symptoms and may cause a significant amount of confusion if these individuals are mixed together in a group. Schinazi [4] further discusses the problems of classification based on visual acuity. He considers two individuals, one with RP the other with myopia who happen to have the same visual acuity. The individual with RP however, also has poor night vision. While still impaired due to a low visual acuity and a restricted visual field, performance in a particular spatial task will probably change in situations with different lighting levels. How can this individual be classified in the same group with someone with myopia whose performance is not strictly dependent on lighting?

The total or partial absence of vision cannot fully account for differences in performance in different spatial tasks. Such strict causality is theoretically sterile and does not recognize the growing evidence on the spatial abilities of the blind and visually impaired [5]. The development of spatial abilities is also mediated through interaction and experience with the environment and culture. In this manner, while individuals in a particular group may have a similar medical, functional or clinical diagnosis (or any combination of these) it does not mean that they are entirely homogeneous. Comparing between groups can be beneficial if the researcher is capable of controlling for a certain amount of cohesion within each specific group. This method is better suited in situations when there are large and clear differences between groups. When differences are small and hard to identify a differential or individual differences approach may be more suitable.

2. The individual differences approach: The importance of subject description and the distinction between ability and present competence

The individual differences approach takes into account the influence different factors can have on the specific individual [6]. It shifts the focus form the actual effect to the possible nature and causes of its presence. Research in visual impairment and blindness is filled with contradictions and interpretations and it is not uncommon to find similar studies with conflicting results. Many of these discrepancies however, can often be attributed to the fact that researchers were working with samples that were not equivalent. In this manner, a crucial step in the analysis consists of detailed description of the characteristics of the group and of each subject as an individual. It is to these descriptions that researchers should refer back when trying to explain the presence of an effect.

Researchers have employed a variety of techniques to assess the extent of the cognitive map knowledge of sighted individuals. Many of these techniques have been adapted for the collection and analysis of data on the content and accuracy of mental representations of space by the blind and visually impaired. Mutually supportive techniques are necessary to account for the heterogeneity of experience and skills between participants and are essential for the interpretation, application and generalization of results. Experiments that assess performance on the basis of one task can generate false conclusions given that performance may be a consequence of a testing artefact rather than ability. A consideration of methodological problems is beyond the scope of this paper. It should be noted however, that both the characteristic of the experiment (difference between spatial memory and spatial inference tasks, size and layout of the space, size of the group, number of elements to process, nature of the response, past experience) and that of the individual (type and age of onset impairment, level of education and intelligence, level of orientation and mobility, affective development) will affect overall performance. In this manner, it is important to recognize and differentiate between ability (possessing the quality to perform) and present competence (the actual performance). Consider a model-building task where an individual is asked to tactually construct a model by positioning different cardboard pieces in order to reproduce the spatial position and relation between different elements in an environment. An individual who fails to perform well in this task can sometimes be labelled as having poor cognitive mapping (at least configurational knowledge) abilities. Here, it is important to consider that this sub-performance may be the result of present competence. This individual may have a perfectly accurate representation of the environment but the actual externalization of the representation is hampered by the task the individual is asked to complete. In this case, a model construction task where successful completion requires not only an accurate representation of the environment but also fine motor skills.

---

[2] Retinitis pigmentosa (RP) is a group of hereditary (pattern varies) diseases that result in the degeneration of the retina. As a progressive disease it begins with damage and destruction to the rods in the mid periphery, gradually advancing to the macula. Tunnel vision, poor night vision and myopia are observed in most cases.

3. Individual differences and the shortcomings of null hypothesis significance testing (NHST)

In the final chapter of his book Blindness and early childhood development David Warren expresses his disappointment regarding the quality of past research on visual impairment [4]. This dissatisfaction stems mainly from methodological and analytical weaknesses that fail to account for the heterogeneity of the population. Researchers often overlook the need for detailed descriptions many times the space and attention forfeited in exchange for statistical significance testing. The problems related to NHST are not new. In fact, they have been around for so long that one cannot help but feel ashamed that for the past two decades a considerable amount of research has been conducted under such an inadequate and insufficient system. The wide availability of computers coupled with user-friendly statistical packages has fuelled a bogus revolution whose main order it to test for significance. This over reliance on statistical testing rather logical reasoning is particularly problematic when used to explain human behaviour. Many of tests commonly used (Student's t-test, ANOVA) rely on group averages and are based on the assumption of a normal distribution. An analysis solely based on group means is unrealistic and restrictive. Significance testing does not tell us whether differences actually matter nor provide any explanations regarding an effect.

Results from statistical testing can be somewhat deceptive, at times giving the illusion of objectivity. As Cohen [7] has impeccably noted, what researchers would like to know is how likely there are differences in the population given the data. Unfortunately, all that NHST can provide is information on how likely is the data given the assumption that there are no differences in the population. Kline [8] points to three false conclusions that can derive from this misinterpretation: 1-The false belief that the p value is the numerical index of the magnitude of an effect. The p value is only a probability statement – the measured probability of a difference occurring by chance given that the null hypothesis is true. Significance levels are highly dependent on sample sizes to the point that increasing the sample size will almost always lead to statistical significance; 2-The false belief that rejecting the null hypothesis automatically asserts the truth (proves) of the alternative hypothesis. Rejecting the null hypothesis does not imply any causal relation. The alternative hypothesis is only one of many possible hypotheses for rejecting the null hypothesis; 3-The wrong conclusion that if you do not reach significance at least at the (.05) level than the study is a failure. As noted by Carl Sagan, "the absence of evidence is not evidence of absence" [9]. Type II errors can occur when statistical power is low or the overall research design and methods are inaccurate. Failure to reject the null hypothesis can be an important part of the research as it forces the researcher to take a closer look at the data.

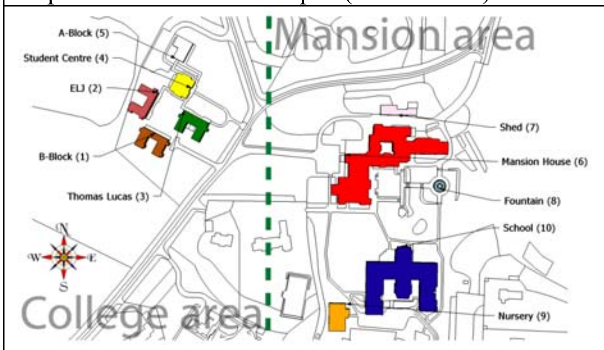
A variety of methods such as effect size measures, confidence intervals that combined with graphical tools can be used as a complement or replacement to NHST. One of the main shortcomings of NHST is that it can only identify the direction of a difference (A > or < B) and does not provide any information regarding its size and amount. The effect size is basically the standardized mean between different groups, a measure of the magnitude of the impact of the independent variable on the dependent variable and can be interpreted as the "true" measure of the significance of a difference. Contrary to p values effect size indicators have the advantage of not being sensitive to sample sizes. More important, they can provide important information on the practical (clinical) significance of a difference (an effect can be statistically significant and mathematically real but too small to be important). Two things should be noted: First, there are no exact guidelines as to what indicates a small or large effect. Some researchers [7, 10] have written substantially on the matter and provide some aids for their interpretation. Second, effect sizes are descriptive and not inferential. They are descriptive of the sample data and offer no information for the rest of the population. Effect sizes should be interpreted with care, as they are highly dependent on the situation. It is up to the investigator to become acquainted with the different characteristics of the data in order to understand what constitutes a small or large effect. Confidence intervals are another useful tool for the interpretation of results and an appealing alternative to NHST given that they can also provide information as to whether the difference between two means is statistically significant. Finally box plots (a visual display of the five number summary) are particularly useful when viewed side-by-side and used to compare data from different groups.

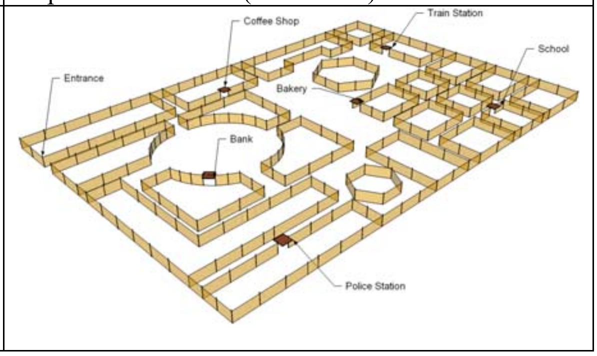
4. Avoiding the statistical deception: Evidence from two studies on the content, accuracy and utility of mental representations

Two experiments were conducted with students from Dorton College at the Royal London Society for the Blind (RLSB). The first experiment examined the content and accuracy of mental representations of a well-known environment. Participants walked a route around the RLSB campus and learned the position of ten buildings and structures. They were then asked to make heading judgements, estimate distances, complete a spatial cued model and sequentially visit a specific set of locations. The second experiment considered the wayfinding strategies and spatial coding heuristics used to explore a complex novel environment. Students were asked to explore a maze and learn the locations of six different places. Their search patterns were recorded and analyzed using Geographic Information Systems (GIS) software. Students were tested with the same methods used in the first experiment and their performance was related to the type and frequency of strategies used. Taking into consideration the previous discussion on the heterogeneous nature of visual impairment and the resulting phenomenological construction of space and reality, data for the two experiments was complemented with a mobility questionnaire, a low vision quality of life questionnaire, a literacy and numeracy assessment as well as ethnographic material collected by the author during the two years spent working and living at the RLSB. Figure 1 below presents the layout for the two experiments.

Fig. 1. Layout for two experiments

Experiment 1: RLSB Campus (10 locations)	Experiment 2: Maze (6 locations)





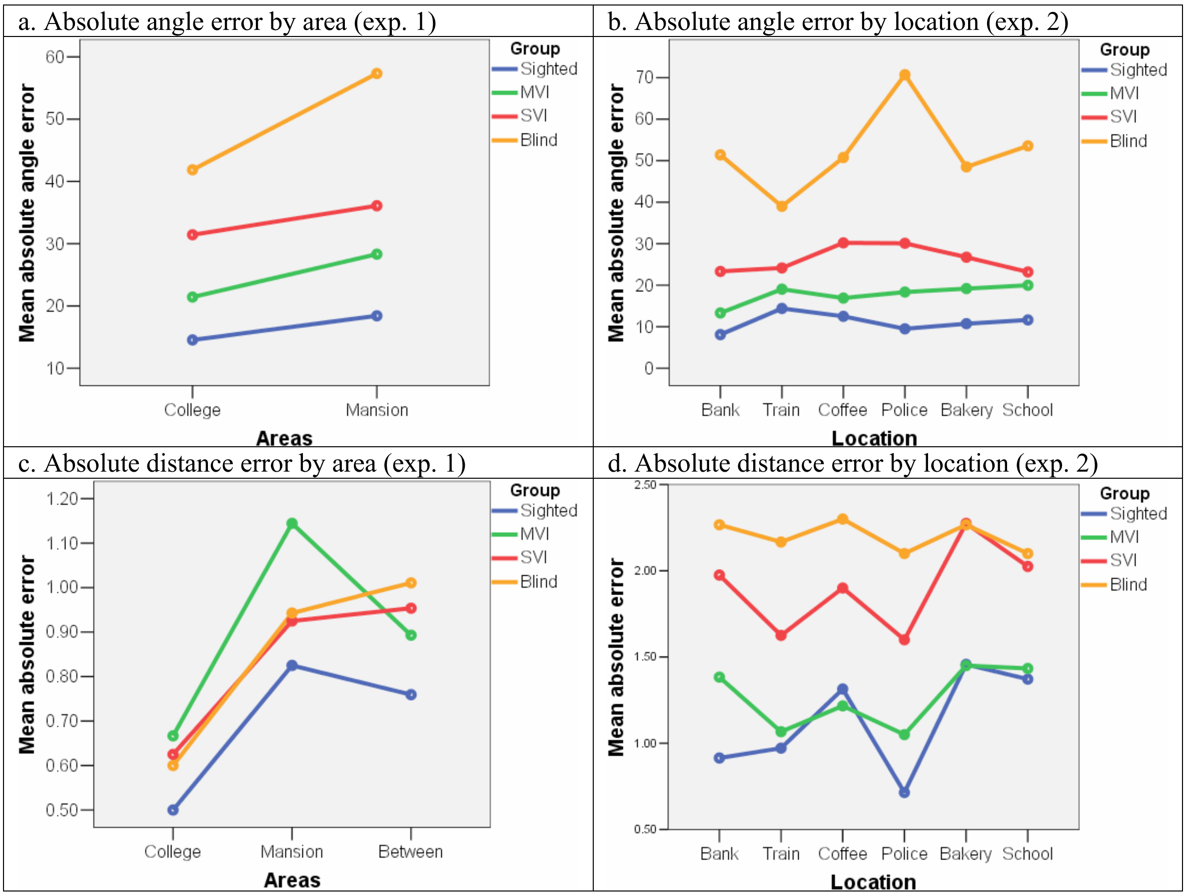
Twenty-nine subjects participated in the first experiment and thirty-three in the second experiment. Subjects were classified into four groups[3] depending on the severity of their visual impairment. The groups are: Sighted, mild to moderate visual impairment (MVI), severe to profound visual impairment (SVI) and blind. For the heading task, subjects were given a digital compass, brought back to each of the locations visited and asked to make a series of pointing judgments. In order to facilitate the analysis of data for the first experiment the ten locations visited were divided between two areas. This division is the result of the distance between two clusters of buildings (and structures) and the academic or administrative functions they serve. A repeated measures analysis of variance (ANOVA) was used to analyze data from the heading judgment task for the two experiments. Figures 2a and 2b present the results of this analysis.

Results from the between group test revealed a significant effect of vision in angle estimation for the first experiment ($F_{3,25} = 7.534$, $p = .001$) and for the second experiment ($F_{3,28} = 11.018$, $p < .001$). A Tukey Honestly Significant Difference (Tukey HSD) post hoc test indicated a significant difference in means between the sighted and the blind (MD = -33.11) and between the MVI and the blind group (MD = -24.73) for the first experiment and between the sighted and the blind (MD = -41.18), the MVI and the blind (MD = -34.53) and the SVI and the blind group (MD = -26.03) for the second experiment. This difference can be clearly seen in the graph where the lines for the sighted and the MVI groups (first experiment) and for the sighted, MVI and SVI groups (second experiment) are at a distance from that of the blind. As expected, the within subject test revealed a significant area effect ($F_{1,25} = 18.007$, $p < .001$) indicating that groups were more accurate when pointing to locations at the college area when compared to the mansion area. The better accuracy when pointing to locations at the

[3] This grouping is based on the International Council of Ophthalmology's (ICOPH, 2006) *multiple ranges of vision classification.*

college may be attributed to past experience and level of familiarity with the area as the college is where the vast majority of the participants live, work or study The interaction between area and group was not significant (lines in the graph are relatively parallel). For the second experiment the within subject test revealed a significant location effect ($F_{5,140}$ = 2,871, p = .017) meaning that the accuracy of pointing estimates changed depending on the location. The interaction between location and group is also significant ($F_{15,140}$ = 2.253, p = .007) meaning that that the accuracy of the pointing estimates changed depending on the location but changed in different ways. Looking at graph 2b we notice that the lines between the groups are not parallel (this is particularly the case for the blind group).

Fig. 2. Results for two experiments on angle and distance estimation tasks



Using Cohen's guidelines, the calculation of *effect size* further supports the presence of a linear relationship between the degree of visual impairment and performance. Table 1 presents the results of a series of t-tests along with the *effect size* for each group combination in the first experiment. According to Cohen's interpretation, the effect size of the two statistically significant means is considered "huge". Here we note that the mean absolute angle error for the blind (50.16º) is twice as large as the MVI (25.10°) and three times larger when compared to the sighted group (16.61º). Regarding the other mean differences there is an actual progression between medium, larger, and very large effect size and the degree of visual impairment.

At first glance, these results may seem somewhat robust and this has led some researchers to discuss the spatial abilities of the blind and visually impaired (sometimes solely!) in respect to their ability to make accurate pointing estimations. However, one should be careful on drawing strict conclusions from one specific test. Although all subjects went through a training period in order get acquainted with the apparatus (digital compass) and testing procedure, pointing to objects at a distance is not an action commonly undertaken by visually impaired and particularly by blind individuals. Here it is important to differentiate between ability and present competence. A study that aims to assess the spatial abilities of the blind and visually impaired but solely tests individuals in regard to pointing accuracy does not allow for such distinction. The fact that blind and visually impaired people are not accustomed to point and will have a greater chance of performing at a lower level (greater absolute angle error) does not mean that the content and accuracy of their mental representation is substandard. Pointing tasks do not allow for a correct externalization of a representation as these groups may have different strategies to collect and process distal information. In this manner, their weak performance may be related to a testing artefact not actual spatial ability.

Table 1. Effect size for angle estimation task (experiment 1)

Group	Mean difference	Test statistic	Effect size Cohen's (d)
Sighted vs. MVI	-8.49	(t = -1.517, df = 12, p = .155)	0.91 (large effect)
Sighted vs. SVI	-17.31	(t = -2.108, df = 11, p = .059)	1.31 (v. large effect)
Sighed vs. blind*	-33.54	(t = -5.436, df = 10, p = .000)	3.49 (huge effect)
MVI vs. SVI	-8.82	(t = -1.241, df = 15, p = .234)	0.64 (medium effect)
MVI vs. Blind*	-25.06	(t = -4.118, df = 14, p = .001)	2.22 (huge effect)
SVI vs. Blind	-16.24	(t = -2.026, df = 13, p = .064)	1.13 (very large effect)

***Significant at the .05 level (equal variances assumed)**

For the distance estimation task subjects were given different sets of three locations (a lambda 4 balanced incomplete block design was used to generate the triad questionnaire for the first experiment) and asked to estimate which two locations are the closest together and which two locations are the furthest apart. An error score was created relative to an ideal[4] Euclidean (straight line distances) and functional (route distances) participant. For the purpose of this paper, error scores are presented only in relation to the ideal Euclidean participant. In order to facilitate the analysis, locations in the first experiment were separated into three areas (college area, mansion area and between areas). A repeated measures analysis of variance was used to analyze data from the distance estimation task for the two experiments. Figures 2c and 2d present the results of this analysis.

Results from the between group test revealed no significant effect of vision in distance estimation for the first experiment. A significant effect of vision was found for the second experiment ($F_{3,29} = 4.869$, $p < .007$). A Tukey HSD post hoc test indicated a significant difference in means between the sighted and the blind (MD = -1.08) and between the MVI and the blind (MD = -.933). This difference can be clearly seen in graph 2d where the lines for the sighted and MVI group are at a considerable distance from that of the blind. For the first experiment the within subject test revealed a significant area effect ($F_{2,48} = 19.031$, $p < .001$) indicating that the accuracy of distance estimations varied between areas. All groups however, were more accurate when estimating the distance of locations situated within the college. The interaction between area and group was not significant. The same was the case for the second experiment where the within subject test revealed a significant location effect ($F_{5,145} = 3.785$, $p = .003$) indicating that accuracy of distance estimation varied between locations. The interaction between location and group was not significant.

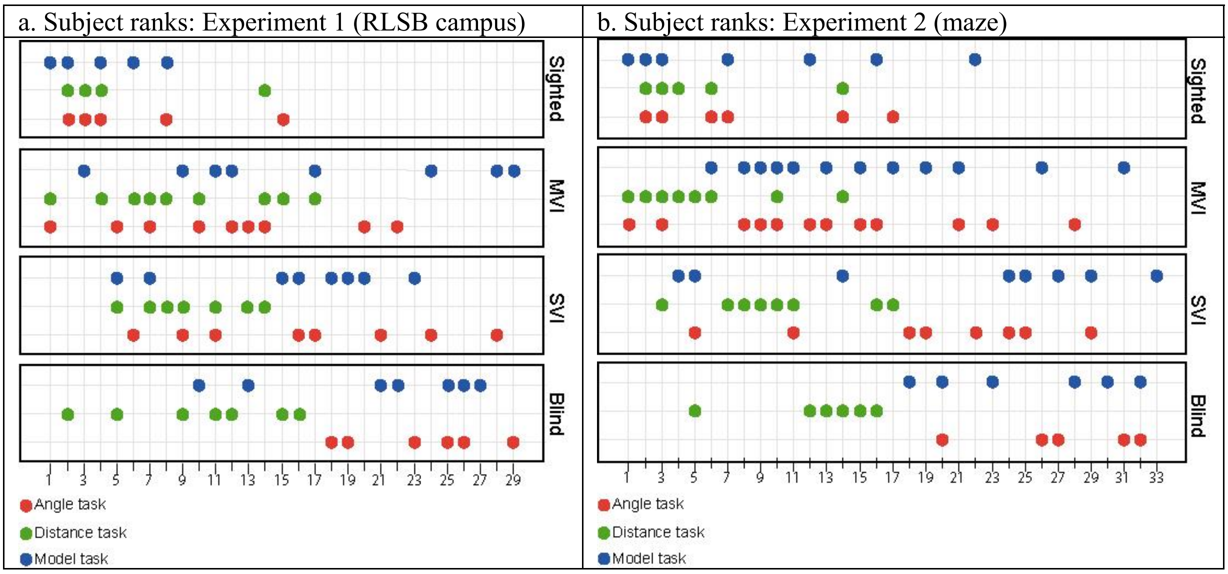
The ANOVA results for the first experiment point to two different trends. Both the sighted and MVI groups were more accurate when estimating distances between locations they were familiar with (locations in the college area) and those between areas (when the actual distance between locations is very large). This was not the case for the blind and visually impaired groups who were more accurate when estimating distances in the college and the mansion area of the campus. For the second experiment all groups tended to behave in a similar fashion, with the blind an SVI groups making less accurate estimations than the sighted and MVI groups. The fact that there are no significant differences between the groups in the first experiment but these are present in the second is consistent with the constructive effect familiarity (the maze was a completely novel environment) can play in the formation of mental representations.

In order to further investigate the results from the ANOVA the performance of each subject was ranked. Figure 3a and 3b presents the ranking distribution of the three tasks for the two experiments (lower ranks indicating a better performance). As mentioned before, differences that are not statistically significant should not be discarded as they force the researcher to take a closer look at the data.

---

[4] The ideal participant is someone who completes the triad questionnaire without any mistakes. Responses from the ideal candidate were generated in GIS (Geographic Information Systems) using an actual map of the space.

Fig. 3. Subjects ranked by performance for three spatial tasks



Looking at the rank standing in this particular task (green dots) for the first experiment we notice the presence of good and bad performers in all four groups. In fact, the best performer (BP) and worst performer (WP) in the distance estimation task belong to the MVI group. For the second experiment there exists a bigger cohesion (dots closer to each other) among the individuals in each group – this is supported by the statistically significant mean differences between the sighted and the blind and the MVI and the blind. A closer look at the characteristics of the BP and WP in the first experiment can provide some possible explanations for the disparity in performance. Table 2 presents some characteristics of these two performers.

Table 2. Best and worst performer characteristics for distance estimation task

Subject	Age	Sex	Visual acuity	Visual field	Colour vision	Visual condition	Age of onset	Other disability	Mobility (years)	Literacy & Numeracy
BP	19	M	0.1666	Constricted periphery	Failed Ishihara	Brain tumor	17	No	1	Lit. 56/72 Num. 46/50
WP	21	M	0.6666	Constricted periphery	OK	Divergent squint & nystagmus	Birth	No	4.5	FSA Level C

At first glance apart form the age of onset impairment[5], the characteristics of the two participants does not seem to provide enough evidence to explain for the difference in performance. Both are males, arrived at the college at about the same time, and have a constricted field of vision. To further complicate matters, the worst performer has a stronger visual acuity and more years of mobility training. However, during the testing phase the author noted that the WP had considerable problems with his short-term memory. In fact, a look at the subject's file and medical history indicated that, although never formally diagnosed, the subject had shown considerable problems with his memory both at college and at home. This is consistent with his scores in the literacy and numeracy assessment, which are well below that of the best performer. The fact that the distance questionnaire was verbally presented to each of the participants and these were required to retain the name of each location in their memory and workout which two of the three locations were the closest together and which were the furthest apart seems to have capitalized on this weakness. In this manner, this subject's inferior performance may be more related to a mental operation problem made difficult due to a short-term memory issues than to the actual ability to represent space. Considering that this subject's performance is in the middle ranks in the angle estimation task further supports this view. Given the familiarity with the campus, pointing responses required both the manipulation of short and long memory stores.

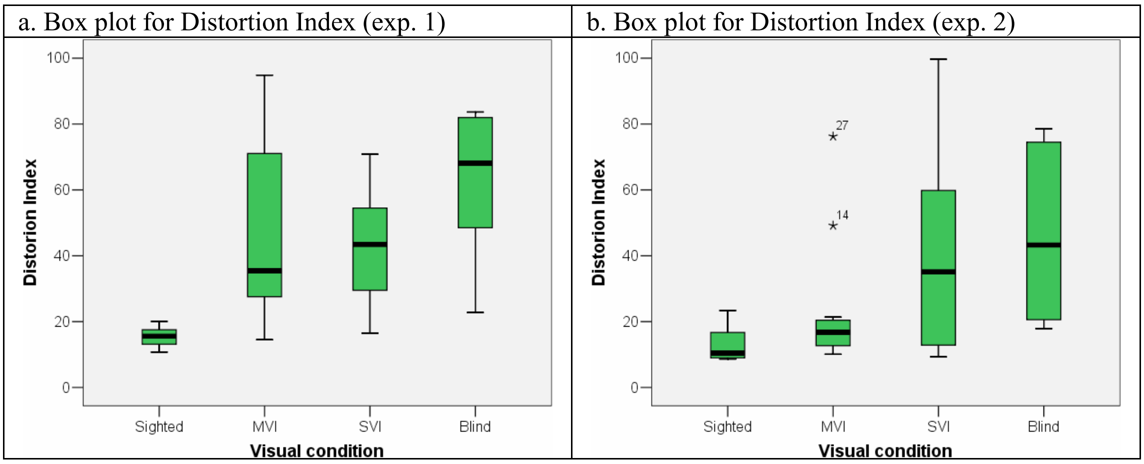
The last task consisted of completing a cued model of the campus. Three-dimensional card pieces (magnets for the second experiment) representing scaled versions of the locations visited (or explored)

---

[5] Here it is important to remember that both subject belonged to the MVI and made good use of their functional vision.

were created. Three of these were fixed on their real cartographic location on a gridded (1cm X 1cm) magnetic white board (47 cm X 39cm). Subjects were asked to position the remaining pieces in relation the fixed ones. The position of the estimated locations was calculated using a coordinate type system (x, y) provided by the actual grid. Results were analyzed using bidimensional regression [11].[6] Bidimensional regression statistically calculates the degree of association (r^2) between two configurations of related coordinate data. Like in a normal regression the r^2 value ranges from zero to one, with one representing a perfect fit. Results from bidimensional regression can also be used to calculate a distortion index (DI), which is basically a measure (%) of the overall distortion of the representation, a standardized measure of relative error. For the purpose of this paper we will concentrate on the distortion index.

Fig. 4. Distortion index box plot for two experiments



The box plots for the two experiments (figures 4a & 4b) indicate that in both cases the sighted was the best performing group. Looking back to the ranking diagram we notice that the top two performers for both experiments belong to this group. More important however, is the fact that good and bad performers are present in the MVI, SVI and blind groups. If we discard the sighted group we notice that for the first experiment the MVI group houses the best and worst performers (look at the low and high whiskers). The same is true for the SVI group in the second experiment. In the first experiment the distribution for the MVI is particularly skewed to the right (mean is larger than the median) pointing to a considerable presence of bad performers. There are two extreme values (subjects 14 & 27) for the MVI group in the second experiment. In regards to the blind group the distribution is slightly skewed to the left (stronger presence of good performers) for the first experiment and slightly skewed to the left in the second experiment. Table 1 below presents the results of the model task (r^2 & DI) along with the minimum and maximum value for each group.

Table 3. Degree of fit (r^2) and distortion index (DI) for model task

Experiment #	Group (n)	r^2 (SD)	DI (SD)	Min (DI)	Max (DI)
1	Sighted (5)	.98 (.01)	15.40 (3.64)	10.74	20.05
1	MVI (9)	.68 (.36)	48.55 (30.72)	14.51	94.85
1	SVI (8)	.78 (.16)	42.78 (18.61)	16.45	70.82
1	Blind (7)	.56 (.27)	62.22 (24.41)	22.75	83.65
2	Sighted (7)	.97 (.02)	13.41 (5.66)	8.65	23.37
2	MVI (12)	.91 (.17)	23.18 (19.64)	10.09	23.37
2	SVI (8)	.75 (.33)	40.57 (31.38)	9.30	99.66
2	Blind (6)	.72 (.27)	46.34 (27.72)	17.90	78.56

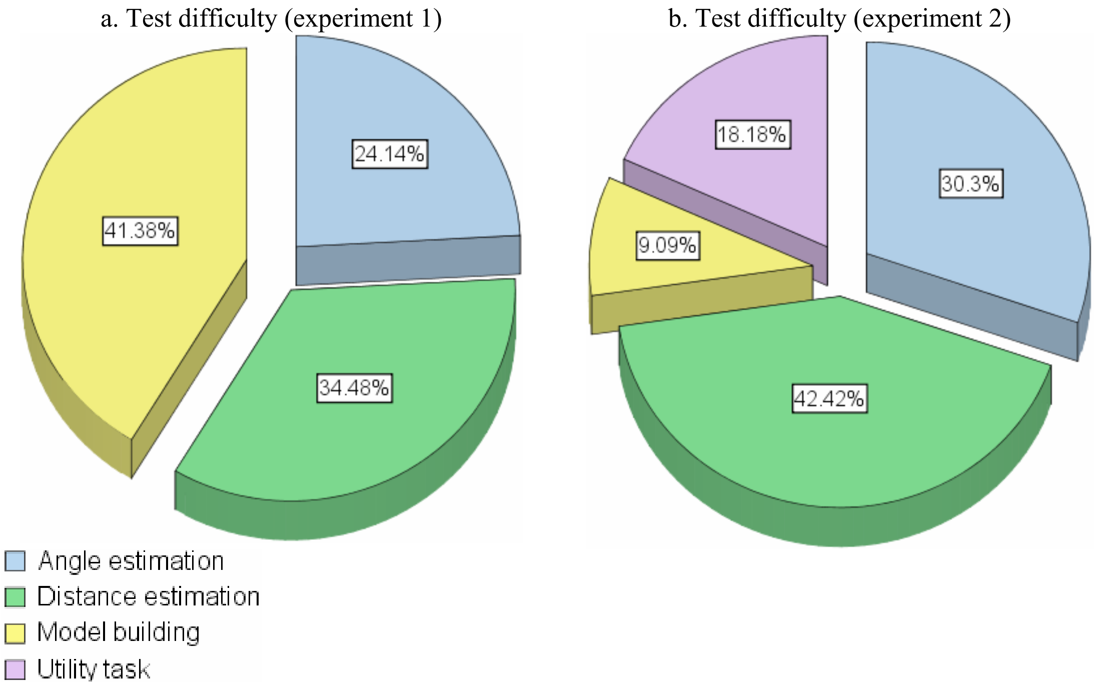
Considering the information gathered through the box plots it was no surprise that Levene's test of homogeneity of variance was violated for the first (.015) and for the second experiment (.014). The original distortion index data was converted to logarithms, squeezing all the data points closer together, to correct for the assumption in an ANOVA that variance is the same across all the data. After the

[6] See Kitchin & Blades "The cognition of geographic space" for a detailed account of this method.

logarithmic conversion Levene's test for the homogeneity of variance was not violated for the first (.246) and second (.131) experiment. Results from the ANOVA revealed a significant effect of vision in the construction of a cued model for the first ($F_{3,25} = 6.397$, $p = .002$) and the second ($F_{3,29} = 3.965$, $p = .017$) experiment. For the first experiment a Tukey HSD test pointed to significant differences between the sighted and the MVI ($p = .013$), sighted and the SVI ($p = .022$) and the sighted and the blind group ($p = .001$). For the second experiment a significant difference existed only between the sighted and the blind ($p = .022$).

As a final piece of evidence on how performance can vary depending on the actual task, participants were asked, in a debriefing questionnaire, which of the tasks they thought to be the hardest. Figures 5a and 5b present the participants perception of task[7] difficulty for the two experiments.

Fig. 5. Participants' perception of test difficulty



In the first experiment, the model-building task (41.38%) was judged to be the hardest test to complete, followed by the distances estimation (34.48%) and the angle estimation tasks (24.14%). This was not the case for the second experiment where the model-building task (9.09%) was perceived as the easiest. The distance estimation (42.42%) followed by the angle estimation (30.3%) tasks were perceived as the hardest tasks. This discrepancy in results may be attributed to the difference in the size of space between the two experiments. It can be argued that the smaller intra-distances between the locations in the second experiment led to more estimation errors. In order to assess whether the four groups differed in their perception of task difficulty two chi-square tests were performed. Results revealed no significant differences between groups on perception of task difficulty for the first experiment ($\chi^2 = 6.568$, $DF = 6$, $p = .363$) and for the second experiment ($\chi^2 = 10.325$, $DF = 9$, $p = .325$). These results are consistent with the performance results for the three tasks where in many cases the same individual performed at different levels depending on the task.

Conclusion

Using data from the two experiments conducted at the RLSB, the article demonstrated the importance of considering individual differences when analyzing data on the mental representation of space. The need of mutually supportive techniques and detailed participant description is evident from the

---

[7] The second experiment had additional task (utility task), which consisted of reproducing a route in the real environment. This task has not been discussed in this paper but results are still part of the task difficulty question.

variability of performance between and within subjects when asked to complete different tasks. The concepts of ability and present competence were introduced and discussed in relation to past experience and the provision and access of information. Finally, several analytical techniques such as effect size estimates, box plots and the actual visualization of subject ranks were presented as a necessary complements to NHST.

Many renowned researchers (Kevin Lynch and Jean Piaget come to mind) have managed to reach important conclusions without relaying on significance testing. Unfortunately many are still under the illusion that results accompanied by significant *p* values are more robust and a fundamental requirement for publication. NHST can sometimes be a trivial exercise as rejecting the null hypothesis is usually the case of securing a large enough sample. Results that are not statistically significant should also be reported as they force the researcher to pay attention to the array of variables that are brought by each individual and are beyond the control of the experimenter. The individual differences approach combines the logic of the case study technique with the advantages of quantitative methods. There are however, several difficulties with this approach. These are related to the vast array of factors (physical, clinical, environmental) that even if identified can have a different effect on each participant. Nonetheless, it is exactly this complexity that should interest researchers. Good research is not only concerned in identifying an effect (or a difference) but also in uncovering the facts that can explain the reasons behind the difference. Explaining the difference is what will aid professionals in the design of intervening programs that are catered to the group or the specific individuals.

References:

[1] Millar, Susanna. Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children, Oxford Science Publications. Oxford: Oxford University Press, 1994.

[2] Lewis, Vicky, and M. Glyn Collis. "Methodological and Theoretical Issues Associated with the Study of Children with Visual Impairment." In Blindness and Psychological Development in Young Children, edited by Vicky Lewis and M. Glyn Collis, 1-8. Leicester: The British Psychological Society, 1997.

[3] Rosa, Alberto, and Esperanza Ochaíta. Psicología De La Ceguera, Alianza Psicología; 39. Madrid: Alianza Editorial, 1993.

[4] Schinazi, Victor. "Spatial Representation and Low Vision: Two Studies on the Content, Accuracy and Utility of Mental Representations." In International Congress Series: Elsevier, 2005.

[5] Warren, David H. Blindness and Early Childhood Development. 2nd ed. New York, NY: American Foundation for the Blind, 1984.

[6] Lewis, Vicky, and Vicky Lewis. Development and Disability. 2nd ed. Malden, MA: Blackwell Pub., 2003.

[7] Cohen, Jacob. "Things I Have Learned (So Far)." American Psychologist 45, no. 12 (1990): 1304-12.

[8] Kline, Rex. B. Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research. Washington, DC: American Psychological Association, 2004.

[9] Sagan, Carl. The Demon-Haunted World: Science as a Candle in the Dark. 1st ed. New York: Random House, 1995.

[10] Cohen, Jacob. "The Earth Is Round (P <.05)." American Psychologist 49, no. 12 (1994): 997-1003.

[11] Kitchin, R. M., and Mark Blades. The Cognition of Geographic Space. New York: I.B. Tauris & Co Ltd, 2002.