Running head: SUBJECTIVE SIGNIFICANCE JUDGMENTS

Subjective significance judgments

Dror Lev and David Leiser

Ben-Gurion University, Israel

Corresponding author:

David Leiser

Department of Behavioral Sciences

Ben Gurion University

84105 Beer Sheva

Israel

Fax: + 972 2 5618841
Phone: + 972 54 4979225

dleiser@bgu.ac.il

Subjective significance judgments

Abstract

Observers were presented with random spatial distributions of dots, inside and outside a small highlighted region, and required to judge to what extent their density was significantly different. Experiment 1 examined whether the absolute density of the target area affects significance judgments, and evaluated the size of this bias compared to that of the normative statistical significance, using regression analysis; the regression weight of statistical significance was about four times larger than the absolute target value. Experiment 2 investigated the effect of increasing the size of the cells, which dilutes the dots in a larger area without affecting the statistical task. This factor produced a small residual effect, suggesting the involvement of a perceptual mechanism in statistical processing. The third experiment found that statistically explicit instructions both increases the influence of the statistical information and reduces the absolute density bias found in Experiment 1.

Keywords: randomness, significance, naïve statistics, classification, categorization

Subjective significance judgments

A fundamental requirement of any perceptual system is the ability to distinguish significant objects or events from random “noise” in the environment. This task is not always done well. Behavioral economists, for instance, have pointed out that human perception of randomness is inaccurate, and this has significant effects on economic behavior (Altmann & Burns, 2005; Black, 1986). While decisions may have to be dichotomous (sell or keep a share), they are underlain by a judgment of a continuous quantity: how significantly is the putative event different from the background noise. The present work will explore *subjective* judgments of significance level.

Prior to our study, observers were not asked to evaluate the significance of deviation from randomness. This is surprising, since the question has such obvious relevance to real-life issues. For instance, the significance of the deviation from expected randomness plays a vital role in the epidemiological analysis of geographic clustering. A troubling problem in public health occurs when a community notices that an unusual number of its residents are stricken with, say, cancer. The question then arises of whether this number constitutes a mere statistical fluctuation from the expected value, or whether something increases the risk of cancer there, and something may be done about it (Anto & Cullinan, 2001; Gawande, 1999; Siegrist, Cvetkovich, & Gutscher, 2001; Thun & Sinks, 2004). Epidemiologists have noted the "tendency of the human mind to identify patterns (and causes) rather than randomness, and lack of social trust in public health experts" (Siegrist et al., 2001). When a community expresses alarm, it makes a judgment: the numbers in their locality, compared to the background prevalence, has reached some statistical significance level, though in such epidemiological cases, their judgment may be heavily biased towards avoiding misses at the cost of increased false alarms. This literature relied on the psychological literature on *subjective randomness*.

Past research on subjective randomness has trod a narrow path, relying almost exclusively on experimental paradigms where participants had to identify or produce strings of characters that they considered to be random. Previous studies also mostly focused on one specific aspect of these stimuli, departure from random alternation and the corresponding occurrences of longer and shorter runs. For example, in his exhaustive review of the literature entitled "*The production and perception of randomness*", Nickerson (2002) restricted his discussion to such cases, as had done Bar-Hillel and Wagenaar (1991) before him. In these studies, the objective (that is, statistical) properties of strings considered as random by the observers are taken to express "subjective randomness". The main finding, observed with both production and judgment tasks, is called the *negative recency effect*. This refers to a bias in the perceived probability of alternation between consecutive characters in a string (Nickerson, 2002). For example, for strings of binary characters and sequential independency where the expected probability of alternation is 0.5, such as with a fair coin, observers use 0.6 as the mark of a random string (Falk, 1975; Falk & Konold, 1997).

The issue of significance level is also related to the study of categorical classification. In general, classification studies are concerned with the processes that underlie the ascription of objects to categories. The most common modeling approach is based on similarities between object and category (Pothos, 2005; Sakamoto, Love, & Jones, 2006). Nevertheless, there is evidence that similarity alone can not explain the entire range of classification phenomena. Specifically, in the domain of perceptual classification it has long been shown that statistical properties of the categories, such as category variability, play a role in classification. This claim can be traced back to Posner & Keele's (1968) classic study of category learning. Rips (1989) is regarded the first to demonstrate this in a non-learning study with real-life categories. Rips' was also the clearly separate the two dimensions: similarity and statistical variables (category variability). This distinction is at the heart of

current studies of the statistical aspects of classification (Cohen, Nosofsky, & Zaki, 2001; Sakamoto, Love, & Jones, 2006) whose consistent findings indicate "… that humans develop distributional knowledge for categories, which they use when making category judgments." (Sakamoto, Love, & Jones, 2006).

The common statistical model for a classification task with two categories is Signal Detection Theory (Green & Swets, 1966), later extended by Ashby & Townsend (1986) to tasks involving more than two categories. This approach requires statistical knowledge of all categories, at a minimum their central tendency and variance. When the statistical properties of only one category is known, the question reduces to whether an object belongs to this category, and the appropriate statistical tool is significance level. Whereas researchers of perceptual classification demonstrated human sensitivity to the statistical variability of categories, they did not do so for the single-category case, hence they never evaluated humans as judges of statistical significance.

The present study introduces an experimental paradigm to study subjective significance level, i.e. how people judge whether a given departure from a background distribution is significant, in the statistical sense. The subjective randomness literature predicts the existence of biases, while the classification literature predicts adherence to the statistical information. In this study we independently manipulate biasing factors and the statistical information and report their relative influence.

Suppose a person has to determine whether a given cluster of cancer cases is significantly atypical relative to a uniform random distribution of cancer cases. The statistical function of significance level has two parameters: a *target value* (the cluster of cancer cases to be judged), and a reference, *background distribution* (the random distribution of cancer cases). We focus on the biasing influence of the first of these, the *target value*. First, we will study how target value and statistical significance affect observers'

judgments. We will then examine to what extent perceptual features and changes in instructions affect the relative reliance on them.

Experiment 1: The impact of statistical information and of biasing data

The first experiment introduces the general experimental paradigm. It is designed to measure the relative influence of the two factors at the heart of this study: the relative influence of the normatively relevant variable, namely the statistical significance level of the target relative to a specified distribution, and the biasing effect of the target value. This will enable us to determine how well observers fare as "significance evaluators" and how far they fall pray to the biasing factor.

Insert Figure 1 about here

Inspired by the cancer clusters problem, the stimuli are two-dimensional distributions of large numbers of points with a grid of squares superimposed to define the clusters (see Figure 1). The target is one of these clusters.

To calculate the statistical significance level we need a suitable statistical index of significance level for clusters in two-dimensional distributions of large numbers of points, and this is provided by the Poisson distribution. The Poisson distribution gives the likelihood that a given cell containing a certain number of points was sampled from a population of cells with a given average number of points (Feller, 1968). We are not interested in the accuracy of observers' answers, but will analyze the pattern of judgments. If observers behave normatively, they should base their decisions solely on the Poisson distribution, while there should be no effect of the target value beyond its contribution to the statistical significance.

Method

Participants. Thirty-two undergraduates engaged in the experiment, as part of a course requirement. The participants were informed that five of them would be randomly chosen and paid up to 50 NIS (about USD10) according to their performance.

Design. A regression design with two predictors: the value of the target cell and the ratio of that value to the mean of the background ("ratio"). The target value ranged from 20 to 120, and the "ratio" ranged from 0.5 to 1.85. We choose these values so that their combinations cover the full range of the cumulative Poisson distribution (0, 1). Observers judged the degree to which the target cell deviates from the background, which we encoded on a scale of [-15, 15] though the slider itself had no markings. These values constitute the dependent variable.

Materials. We constructed each stimulus by selecting a target cell at random (see Figure 1). We then selected how many dots to place in this cell (target value) from a uniform distribution in the range, also at random. Dots were placed inside the target cell according to a uniform spatial distribution. Next, the "ratio" was selected in the same way. This yielded a density value for the background, and the required number of dots was distributed uniformly over the rest of the frame. The overall size of the picture was 480x480 pixels, each cell was 60x60 pixels, and the square dots were 3x3 pixels.

The cover story mentioned a group of rural plots for sale, each with a different potential for growing crops. Observers judged the potential of a certain plot (marked) among many others (the background cells), according to the concentration of dots inside them, the dots representing plants (see Appendix A).

Procedure. Participants were shown displays such as Figure 1. To minimize perceptual influence, the frame was invisible most of the time. Whenever observers wanted to see which cell was marked, they would press the "display" button and see the frame for half a second. When ready, they indicated, with the slider, to what extent they judged the area marked with the frame to be significantly different from the rest. A central placement of the slider signaled that the target (marked) cell is no different from the rest. They pushed the slider to the right or the left to signal that the target cell was more or less promising than the background.

Results

We computed, for each trial, the statistical significance (Stat-Sig) of the difference in density of the target cell compared to that of the background, using the average number of dots per cell in the background as the parameter of the Poisson distribution. Next, we regressed the observer's judgments on significance and target value. The overall variance explained was $R^2 = 0.42$, and the model was highly significant $F(2, 3357)=1219$, $p<.0001$. The regression coefficients were: significance (Stat-Sig) $\beta_{StS}=0.635$, $p<0.001$ and target value: $\beta_T=0.167$ $p<0.001$, β_T being about quarter of β_{StS}.

Insert Table 1 about here

We computed the regression weights for the individual participants, along with the β_T / β_{StS} ratio (Table 1). As may be seen, the results across participants are also found at the individual level, with a median ratio of 0.12. Judgments are affected by both factors. The strongest influence is β_{StS}, that of the statistical value (Stat-Sig) they were requested to evaluate, but they are influenced by the absolute (target) value too, though to a significantly lesser extent [$t(31)=6.12$, $p<0.001$].

Experiment 2: Perceptual manipulation and instructions elaboration

Experiment 1 showed the biasing effect of the amount of dots plotted in the target cell, over and above its deviation from the background distribution. Changing the amount of dots without changing the size of the cell, as was done in Experiment 1, is a manipulation of density. The biasing effect of the absolute amount of dots may therefore have a perceptual explanation in terms of density considerations. The human perceptual system is known to be sensitive to density changes (spatial frequency) (for a review see Bruce, Green, & Georgeson, 2003). Meyer, Taieb, & Flascher (1997), who studied perception of correlations presented as scatterplots, recommend that "… instead of dealing with the subject domain (e.g., statistics) when trying to understand intuitive judgments, one should analyze the

geometric and perceptual properties of the displays or other information on which estimates are based." The findings in Wiegersma's (1987) study of perceptual influences on the *negative recency effect* support this view, and suggest that judgments of randomness are affected by perceptual processes.

Experiment 2 evaluates the residual impact of a strictly perceptual manipulation. This experiment follows the design of Experiment 1, but we manipulate one additional variable, cell size. For a given number of dots to be plotted, increasing the size of the cell is a straightforward way to decrease the density within the cell, without affecting the Poisson statistic, making it a purely perceptual manipulation. The influence of cell size on subjective significance will be measured over and above those of objective statistical significance level and of target value. This will enable us to identify and assess the importance of a purely perceptual effect on statistical significance judgments.

A second purpose of this experiment is to evaluate whether the explicitness of the instructions matters. To what extent did the performance in Experiment 1 depend on the explicit reference to the need to evaluate the target cell *relatively* to the other cells, as opposed to a mere request to make sure the departure is significant?

Method

Participants. Thirty undergraduates participated in the experiment, as part of a course requirement. The participants were promised that we would pick five of them at random and pay those up to 50 NIS according to their performance.

Design. The design is similar to that of Experiment 1. Besides the two factors manipulated in Experiment 1, target value and "ratio", we also manipulated "cell size". We ran the same experiment on two groups, to which the participants were randomly assigned: one with an impoverished set of instructions, the other with the more explicit instructions we already used

in Experiment 1. Judgments of the degree to which the target cell deviates from the background made up again the dependent variable.

Material and Procedure. These were the same as in Experiment 1, except that cell size was manipulated. The pictures consisted again of a grid of 8x8 cells. The size of the dots remained constant, at 3x3 pixels. The side of the cells ranged from 39 pixels to 66 pixels, so the side of the whole picture ranged from 312 to 528 pixels. Expressed in terms of dots, the side of the cells ranged from 13 to 22 dots.

We used two cover stories: A short one asking the participants to judge the potential of the target cell according to the concentration of dots inside the cells; and a more elaborated cover story, that added the instruction to judge the potential of the target cell in comparison to the background cells. The elaborate story was the one used in Experiment 1.

Results

We ran again a multivariate regression analysis. Using the Poisson statistic, we computed for each trial Stat-Sig, the *significance* of the target cell density relative to the background density, and used this as one of the predictors for the observers' responses, along with target density value, cell size, and instruction explicitness. The overall variance explained was $R^2 = 0.30$, and the model was highly significant $F(3,2996)=436$, $p<.0001$. The regression coefficients were as follows: significance (Stat-Sig): $\beta_{StS}=0.49$, target value: $\beta_T=0.24$, and cell size: $\beta_{CS}=0.07$; all $p<0.001$. β_T is about half the size of β_{StS}, and β_{CS} about 0.14 its size. The instruction explicitness variable was not significant ($\beta_I =0.004$, $t(2999) =0.31$, $p=.07$). We next computed the regression values for the *individual* participants across instructions. The median value for the significance coefficient, Stat-Sig (β_{StS}) is 0.64, the ratio of the coefficients of target value to that of significance (Stat-Sig) $\beta_T / \beta_{StS} = 0.45$ and that of cell size to significance (Stat-Sig), $\beta_{CS} / \beta_{StS} = 0.22$ (Table 2 details the findings).

These results reproduce the findings of Experiment 1 both overall and at level of individuals (compare Table 1).

Insert Tables 2 and 3 about here

Table 3 summarizes the mean coefficients (that of Stat-Sig, of target density, and of cell size) for the two experimental conditions, short and elaborate instructions. We examined the effect of the explicitness of the instructions on the individual coefficients with a mixed two-way ANOVA, taking instruction set (short, elaborate, between subjects) and coefficient (Stat-Sig, target value, and cell size, within subjects) as independent variables. We performed planned comparisons for the three parameters. The specific effect of instructions on the size of the significance (Stat-Sig) coefficient was significant $t(1)=2.49$, $p<0.05$, whereas manipulating the instructions did not affect the *biasing* coefficients at all (the values are identical, and $p>0.95$ for both the target value and the cell size coefficients). The overall differences between the mean weights (across instructions) also proved significant ($F(2, 56)=36$, $MSE=0.03$, $p<0.001$). The effect of Stat-Sig was stronger than the two other factors. In conclusion, the strictly perceptual factor of cell size did have a unique effect, over and above those of the statistical information (Stat-Sig) and that of the amount of dots in the target cell, supporting the claim that subjective judgments of statistical significance are affected by perceptual properties. Further, judges did pay attention to the specific instructions they received and were swayed by them. When explicitly asked to make a relative comparison, they relied more on the relative density of target and background, and this increased the weight of statistical significance (Stat-Sig). This increased reliance on the relative was not accompanied by an absolute reduction of the biasing factors.

The last experiment examines whether drawing attention to conceptual, statistical considerations improves performance.

Experiment 3: Statistical elaboration

Inferential statistics in general and estimates of statistical significance in particular are grounded on the fundamental distinction between the sample examined and the overall population. In Experiments 1 and 2, this distinction was not explicitly mentioned. The present experiment stresses it. We will examine whether this emphasis improves sensitivity to the statistical properties of the display and whether it diminishes the target value bias.

Method

The method was the same as for Experiment 1, except that the instructions emphasized the sample-population distinction (see Appendix A). Thirty undergraduates participated in this experiment as part of a course requirement.

Results

We again calculated for each trial the significance of the target cell relative to its background (Stat-Sig), and regressed the observer's judgments on significance (Stat-Sig) and target value. The model was highly significant $F(2,3147)=1961$; $p<.0001$, $R^2 = 0.55$. The regression coefficients were: significance (Stat-Sig) $\beta_{StS}=0.74$ $p<0.001$ and target value: $\beta_T=0.09$ $p<0.001$, β_{StS} being about eight times larger than β_T.

Insert Table 4 about here

We confirmed that the results across participants reflect those of individuals by running regression analyzes for the individual participants. The values, summarized in Table 4, show a pattern similar to that of Exp. 1 (see Table 1), though with larger differences between the coefficients. We tested our hypotheses with a two-way ANOVA with experiment (1-no statistical details vs. 3-statistical details in the instructions) as a between-subject factor and coefficient (Stat-Sig vs. target density) as a within-subject factor. Planned contrast analyzes showed that the more detailed statistical instructions (Experiment 3) increased the

significance (Stat-Sig) coefficient $t(1) = 2.14$, $p<0.05$, while the decrease of the target value coefficient was only marginally significant ($t(1) = 1.96$, $p=0.054$).

The more explicit statistical instructions affected judgments by enhancing the use of statistical features, and possibly also by lessening the perceptual effect.

General Discussion

The present investigation extends the experimental study of randomness perception to judgments of significance, and developed a novel experimental paradigm for the purpose. The first experiment introduced the general paradigm, and measured the unique influence of the two factors at the heart of the study: the statistical significance level of a target value (relative to a specified distribution), and the biasing potential of that target value by itself. Both factors were found to affect judgments, with the impact of the statistical significance being about four times larger than the biasing impact of the target value.

Experiment 2 evaluated the residual impact of a strictly perceptual manipulation, the size of the cells. This factor was found to have a unique effect, over and above the effects of the target value and of the statistical significance. That influence is small, about one-fifth the size of the only normative influence (objective statistical significance), and about half that of the target value. These results support the claim that subjective statistical judgments are guided by perceptual processes.

A second purpose of Experiment 2 was to test the influence of the explicitness of the instructions. The instructions in Experiment 1 mentioned the need to evaluate the target cell *relatively* to the other cells. In Experiment 2 we checked how the three factors (statistical significance, target value, and cell size) are affected by such an explicit direction, as opposed to a mere request to make sure the departure is significant. When instructed to make a relative comparison, observers increased their reliance on the statistical significance.

Nonetheless, the biasing influences weren't reduced. Apparently, those biases cannot be overridden by mere operational instructions.

Experiment 3 studied another, conceptual effect of instructions, this time by stressing the distinction between a sample and the population one wants to infer about. Comparing the results of Experiment 3 with that of Experiment 1 showed that this additional statistical emphasis increases the impact of the statistical properties of the display, while diminishing the target value bias.

Our results extend what could be inferred from classification studies: people are sensitive to the statistical structure of categories and use this statistical information for their judgments. Statistical information is not the sole influence on judgments, and the three experiments show a consistent pattern. The major influence on subjective judgments of significance level was the statistical significance level, as is normatively appropriate, but this effect was accompanied by minor biases (less then half the impact of the statistical information). The biases studied were those of the density of the target cell, and the density of the entire display. These effects showed sensitivity to instructions manipulations.

The importance of telling whether a deviation from a norm is statistically significant is of vast practical importance in many applied domains besides epidemiology: medical imagery, behavioral economics, intelligence, agricultural satellite surveys and others. Our study established that people deal with such problems fairly well with some relatively minor biases.

References

Altmann, E. M., & Burns, B. D. (2005). Streak biases in decision making: data and a memory model. *Cognitive Systems Research, 6*(1), 5-16.

Anto, J. M., & Cullinan, P. (2001). Clusters, classification and epidemiology of interstitial lung diseases: concepts, methods and critical reflections. *European Respiratory Journal, 18* (Suppl. 32), 101S-106S.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*, 154–179.

Bar-Hillel, M., & Wagenaar, W. (1991). The Perception of Randomness. *Advances in Applied Mathematics, 12*, 428-454.

Black, F. (1986). Presidential Address: Noise. *Journal of Finance, 41*, 529-543.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). *Visual perception : physiology, psychology, and ecology* (4th ed.). New York: Psychology Press.

Cohen, A., Nosofsky, R., and Zaki, S. (2001). Category variability, exemplar similarity, and perceptual classification. *Memory and Cognition, 26*, 1165-1175.

Falk, R. (1975). *Perception of randomness.* Unpublished doctoral dissertation (in Hebrew, with English abstract), Hebrew University, Jerusalem, Israel.

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104(2),* 301-318.

Feller, W. (1968). *An introduction to probability theory and its application* (3rd ed.). New York: Wiley.

Gawande, A. (1999, February, 8). The Cancer-Cluster Myth. *The New Yorker,* 34-37.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Meyer, J., Taieb, M., & Flascher, I. (1997). Correlation estimates as perceptual judgments. *Journal of Experimental Psychology: Applied, 3*(1), 3-20.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review, 109*, 330-357.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353-363.

Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral and Brain Sciences, 28*, 1-14.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). New York: Cambridge University Press.

Sakamoto, Y., Love, B. C., & Jones, M. (2006). Tracking variability in learning: Contrasting statistical and similarity-based accounts. In R. Sun and N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Vancouver, Canada: Cognitive Science Society.

Siegrist, M., Cvetkovich, G. T., & Gutscher, H. (2001). Shared Values, Social Trust, and the Perception of Geographic Cancer Clusters. *Risk Analysis, 21*(6), 1047-1054.

Thun, M. J., & Sinks, T. (2004). Understanding Cancer Clusters. *CA: A Cancer Journal for Clinicians, 54*(5), 273-280.

Appendix A - Instructions

Experiment 1 and 2

The government is putting on sale agricultural plots of varying agricultural promise in several regions. You are a farming corporation, and interested in buying some plots. You will be shown aerial photographs of the areas. The areas are divided into square plots, and in each area one of the plots (the one for sale) is marked. Agricultural promise may be estimated by the concentration of dots (that correspond to crops) in each plot. You are to evaluate the plots for their fertility, and decide to what extent you would be willing to pay a premium for the marked plot, relative to the going rate in the area. To what extent is the marked area really more fertile than the others?

Experiment 3

You are a farmer interested in buying more land. You will be shown aerial photographs of various areas. Each dot in the photograph is a plant. Generally speaking, fertile plots produce more plants. However, there can be fluctuations in the yield of plots that are equally fertile. It is known that all the plots on the picture but one belong to the same region, and they are all equally fertile. You are to evaluate the fertility of that one plot (marked in blue) in comparison to the fertility of the region to which the other plots belong.

Author Note

We gratefully acknowledge the contributions of Idit Lev, Keren, Ranit, Dveer, and Ifat. We also thank Joachim Meyer, Yaacov Kareev, and Maya Bar Hillel for helpful discussions. The study was supported by a Kreitman Fellowship to the first author and a seed grant from the Office for Sponsored Research at Ben-Gurion University to the second.

Please address correspondence to David Leiser, Department of Behavioral Sciences, Ben-Gurion University, POB 653, Beersheva 84105, Israel. Email: dleiser@bgu.ac.il.

Table 1

Weights of statistical and biasing predictors – analysis on individual subjects (Exp. 1)

Weight	Mean	Median	Percentile 25	Percentile 75
β_{StS}	0.68	0.78	0.55	0.84
β_{T}	0.18	0.09	0.06	0.21
β_{T} / β_{StS}	1.14	0.12	0.08	0.30

Table 2

Regression weights of statistical and biasing predictors – analysis on individual subjects (Exp.2)

Weight	Mean	Median	Percentile 25	Percentile 75
β_{StS} Stat-Sig	0.53	0.64	0.28	0.72
β_{T} target value	0.25	0.23	0.12	0.31
β_{CS} cell size	0.10	0.12	0.08	0.17
β_{T} / β_{StS}	0.76	0.45	0.23	0.78
β_{CS} / β_{StS}	1.15	0.22	0.13	0.38

Table 3

Mean regression weights of statistical and biasing predictors by instructions-analysis on individual subjects (Exp. 2)

	Instruction	
Weight	Elaborat	Short
Stat-Sig e	0.63	0.43
Target Value	0.25	0.25
Cell Size	0.10	0.10

Table 4

Regression weights of statistical and biasing predictors – analysis on individual subjects (Exp. 3)

Weight	Mean	Median	Percentile 25	Percentile 75
β_{StS} Stat-Sig	0.79	0.80	0.74	0.84
β_T target value	0.10	0.10	0.02	0.15
β_T / β_{StS}	0.14	0.11	0.04	0.20

Figure Captions

Figure 1 - *Trial example*

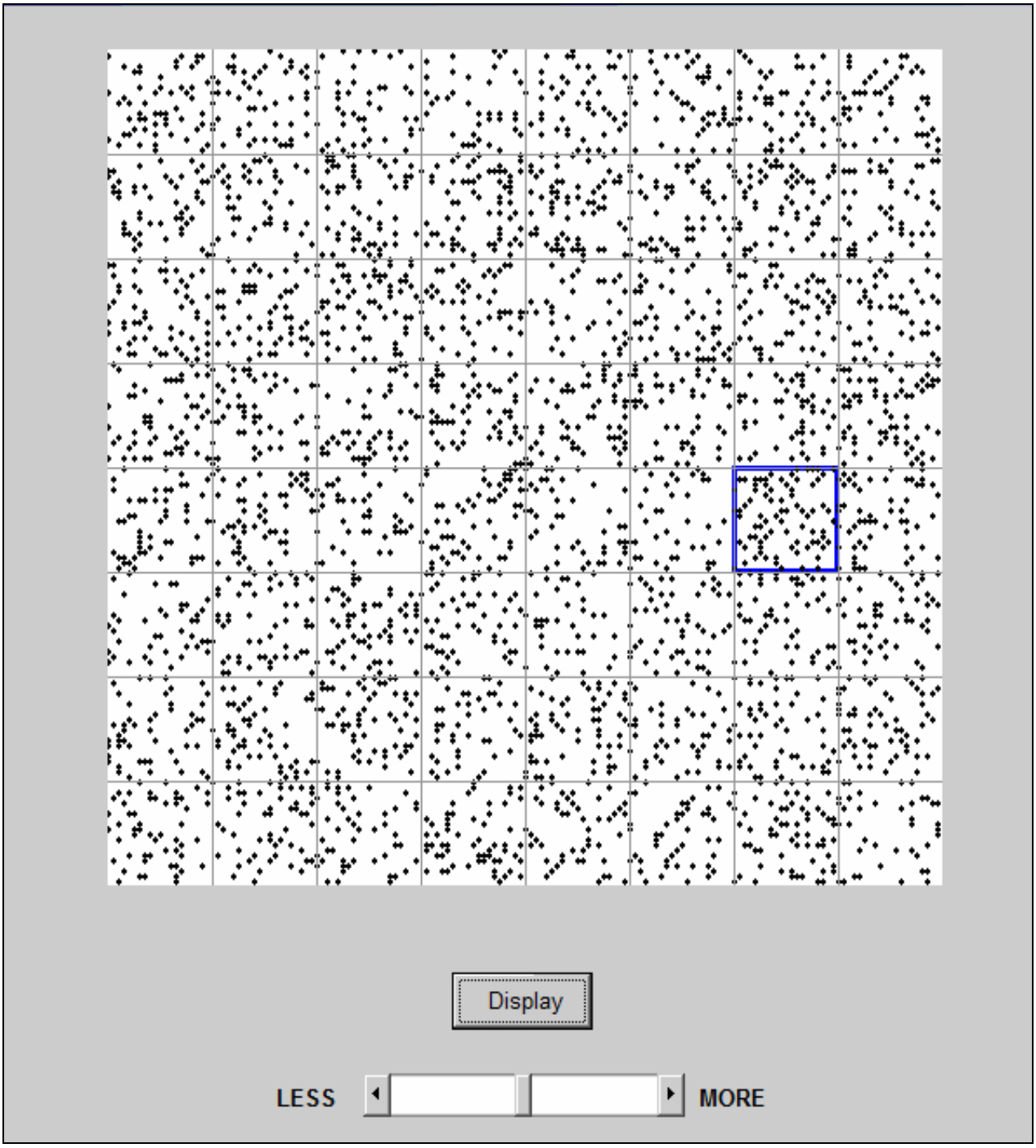

Figure 1