Pre-revised author's version of

(2006) Vickers, D., Lee, M.D., Dry, M., Hughes, P., & McMahon, J.A. "The aesthetic appeal of minimal structures: Judging the attractiveness of solutions to Traveling Salesperson problems". Perception & Psychophysics. 68:1, 32-42.

ABSTRACT

Ormerod and Chronicle (*Perception & Psychophysics*, 1999, *61*, 1227-1238) reported that optimal solutions to Travelling Salesperson Problems were judged as aesthetically more pleasing than poorer solutions and that solutions with fewer internal nodes were rated as more attractive. To test these conclusions, solution regularity and the number of potential intersections were held constant, while solution optimality and the number of internal nodes and the number of nearest neighbours in each solution were varied factorially. A Bayesian analysis of mean ratings of attractiveness showed that the most important single factor was the difference from a benchmark (putatively optimal) solution and that the number of internal nodes made only a minor additional contribution. These results are interpreted as suggesting that aesthetic judgment may be an outcome-oriented metacognitive evaluation of the optimality of a presented structure rather than a process-oriented reflection of the degree to which that structure incorporates spontaneously occurring perceptual principles.

The experience of beauty, of the aesthetic appeal of a painting, a poem, a mathematical proof or a scientific theory, is a pervasive but poorly understood metacognitive aspect of our perceptual and cognitive activity. However, a common thread in much theorising about the basis for aesthetic experience and judgment is the notion of economy, the idea that the judged object represents a parsimonious way of capturing a potentially much greater degree of complexity. In this sense, aesthetic appeal seems to be closely related to the achievement of an optimal solution to a complex combinatorial problem. For example, McMahon (2000) has suggested that what we perceive as aesthetically attractive has to do with "perceptual processes involved with detecting within-object relations in such a way as to cause us to experience … the object … as a solution to the problem of constructing a cohesive form" (p. 34).

From this perspective, the notion of aesthetic appeal is also closely related to the Gestalt notion of figural goodness. For example, Attneave (1982) proposed a speculative realisation of a 'soap bubble' system for implementing the Gestalt law of Prägnanz, which holds that the perceptual organisation of a stimulus will be as 'good' as the prevailing stimulus conditions allow. Attneave's model assumes that the perceptual system operates recursively on pairs of points (or 'dipoles') in such a way as to maximise the coherence among sets of dipoles. Thus, Atteave's model describes an optimisation process, in which the interplay of relations among the dipoles moves the system to an equilibrium or optimum state that is experienced as the 'best' organisation.

In a recent paper, Ormerod and Chronicle (1999) provide some direct evidence consistent with this interpretation of figural goodness. They presented participants with solutions to visually presented versions of the well-known Travelling Salesperson Problem, traditionally referrred to as the Travelling Salesman Problem (TSP). In the planar Euclidean form of the problem, there are n interconnected cities, represented by n nodes, and the task is to construct a pathway that departs from one node to visit each node exactly once, returns to the starting node, and is as short as possible. The number of possible pathways is equal to $(n - 1)!/2$. As a result, when n is small (e.g., 5), there are relatively few pathways to choose from (in this case, 12). However, when n becomes even moderately large (e.g., 25), the number of possible pathways becomes so great that a computer evaluating a million possibilities per second would take almost 10 billion years to evaluate them all (Stein, 1989).

The TSP instances employed by Ormerod and Chronicle contained just 10 nodes in six different configurations. In addition to a standard ("Dantzig") configuration, the remaining five instances contained between one and five 'internal' nodes that were located inside the convex hull bounding the configuration. For each instance, five solutions were constructed in the ranges 0%, 11%-18%, 21%-28%, 32%-37%, and 41%-47% longer than the optimal solution. Participants rated the perceived goodness of each solution, presented as a geometrical figure, on a scale from 1 to 5, where 1 represented a good and 5 a poor figure.

Ormerod and Chronicle found that participants judged optimal solutions as being good figures and progressively less optimal solutions as being progressively less good figures. They also found that participants rated the solutions to problem instances with fewer internal nodes as better figures. They interpreted these results as evidence for the hypothesis that the impressive human ability to arrive at near-optimal solutions to such visually presented optimisation problems may be based on a perceptual organising process in which the optimal pathway corresponds to the figure that is perceived as having the highest degree of figural goodness.

Some support for this view, and also for the identification of figural goodness with aesthetic appeal, is provided by a study by Vickers, Butavicius, Lee and Medvedev (2001). In their first experiment, these authors presented two groups of participants with six different arrays of irregularly positioned dots, each representing the nodes of a TSP instance: two 10-node, two 25-node and two 40-node arrays. Participants in the 'Gestalt' group (G) were instructed to draw a pathway that passed through each of the points just once, that returned to the starting point and that looked "most natural, attractive, or aesthetically pleasing". Participants in the 'Optimisation' group (O) were instructed to draw a pathway that passed through each point just once, that returned to the starting point and that was as short as possible.

Although Vickers et al. (2001) found that there was a relatively small (but significant) difference in the average length of the pathways produced by the two groups, there was also a considerable overlap between the pathways produced by the two groups, with several participants from one group producing exactly the same pathways as those from the other group. This finding is consistent with the conclusion by MacGregor and Ormerod (1996) that the finding of near-optimal solutions to visually presented TSP instances is mediated by spontaneous perceptual organising processes, similar to those underlying the perception of good Gestalten.

At the same time, however, Vickers et al. found that a few participants in group G produced pathways that were much longer than the optimal. In these cases, it seemd that participants were trying to produce maximally symmetric pathways (including pathways characterised by spiral symmetry). These exceptions point to a possible difficulty with the above conclusion drawn by Ormerod and Chronicle regarding the identification of the shortest pathway with the figure judged to have the highest degree of goodness. As the authors themselves make clear, "it is still conceivable that TSP solutions are provided through some other set of cognitive operations, unconnected with a judgment of good figure, and that the relationship with good figure … is merely coincidental" (Ormerod and Chronicle, 1999, p. 1230).

A number of other possibilities, besides solution optimality and number of internal nodes, do indeed suggest themselves as potential factors contributing to the judgments of figural goodness observed by Ormerod and Chronicle. First, the convex hulls for the six figures used in their Experiment 1 (Ormerod & Chronicle, 1999, Fig. 1) describe polygons with progressively decreasing numbers of sides (from 9 to 4, in steps of 1) and with close to perfect regularity. Thus, degree of perceived rotational symmetry is a possible contributing factor. Second, the actual links comprising the optimal solutions shown in their Figure 1 also vary in regularity, as measured by progressive increases in the standard deviation in length of the connecting edges. (From scanned versions of their figures, we estimate the standard deviations to take values proportional to 0.029, 0.039, 0.111, 0.112, 0.141, and 0.236 units for the figures with 1 to 5 internal nodes and the Dantzig figure, respectively.)

Third, the arrays vary with respect to the total number of potential intersections associated with all possible pairs of links between all 10 nodes of each figure. Reliable estimates are difficult to calculate because permissible links (and optimal solutions) are in some cases very sensitive to minute displacements of certain nodes, particularly in the case of the Dantzig figure. However, we estimate that the total number of potential intersections, expressed as a percentage of all permissible links, took values of 29.5%, 25.9%, 21.6%, 19.8%, 19.1%, and 18.7% for figures with 1 to 5 internal nodes and the Dantzig figure, respectively.

Vickers, Lee, Dry and Hughes (2003) found that TSP performance improved as a function of the number of potential intersections in an array and explained this in terms of a hierarchical solution process based on the detection of nearest neighbour links. Since links between nearest neighbours cannot cross, an increase in the number of potential intersections constrains the selection of possible links, thereby reducing response uncertainty and facilitating performance. Thus, this factor also might have contributed to the results of Ormerod and Chronicle (1999).

In addition, there may be other related differences. For example, Ormerod and Chronicle (1999) do not specify whether the solution figures with progressively longer pathways contain any actual intersections. Because the range of solution lengths employed by Ormerod and Chronicle is about twice that which we have encountered in human TSP performance (45%, as opposed to around 22% above optimal), it seems possible that their less optimal solutions might have contained an increasing number of actual intersections.

In addition to theses possibilities, there is a further, plausible alternative hypothesis linking the judged goodness of a TSP solution to objective characteristics of the figure. Specifically, it has been shown that nearest neighbour links are a determining factor in the perception of clusters and motion in random dot patterns

(Dawson, 1991; Vickers, Preiss, & Hughes, submitted; Zahn, 1971), as well as playing a central role in the discrimination of Glass patterns (Caelli, 1981; Vickers, Dry, Lee, & Hughes, submitted). Nearest neighbour links have also been found by Vickers, Bovet, Lee and Hughes (2003), Vickers, Lee et al. (2003) and Vickers, Mayo, Heitmann, Lee, & Hughes (in press) to be important in the solution of visually presented TSP instances as well as in finding solutions for other minimal structures, such as minimum spanning trees (Ahuja, Magnanti, & Orlin, 1993) and generalised Steiner tree problems (Hwang, Richards, & Winter, 1992). Given this suggestive overlap, it seems reasonable to hypothesise that a TSP solution might be judged to have figural goodness according to the extent to which it incorporates naturally perceived structure. That is, solutions that include a greater proportion of nodes linked to their nearest neighbours might be seen as better figures than those with a lesser proportion of such nodes.

Consistent with this hypothesis, it appears that 70% of the links for Ormerod and Chronicle's optimal solution to the figure with just one interior point are mearest neighbour links, while only 50% of those for the optimal solution to the Dantzig figure are nearest neighbours. However, we do not have the coordinates that would allow us to estimate the proportions for the other, non-optimal solutions employed by Ormerod and Chronicle. Accordingly, the following experiment was undertaken to differentiate between the above factors potentially underlying judgments of figural goodness of TSP solutions and to arbitrate between three alternative hypotheses concerning the relationship between these factors and the perceived goodness or aesthetic appeal of such structures. These hypotheses were that aesthetic appeal is based on: (1) the optimality of a solution; (2) the number of internal nodes, or (3) the number of nearest neighbours incorporated in the solution.

In this experiment, the number of potential intersections was held constant, while the number of internal nodes, the optimality of solutions and the number of nodes linked to their nearest neighbours that were incorporated in these solutions were varied in a factorial design. In addition, the regularity of the figures, as measured by the standard deviation in the link lengths, was restricted to a very narrow range and was not allowed to vary sytematically. This was achieved by generating large numbers of random arrays and selecting from them those configurations with the desired characteristics. (This method contrasts with that used by Ormerod and Chronicle to generate their configurations, which consisted of incremental transformations of one initial, highly regular array.)

METHOD

Stimulus generation. In order to obtain expectations for the number of nodes on the convex hull, the number of internal node intersections[1], and the standard deviation of the internode distance, 20,000 random 25-node arrays were generated within the unit square. For these arrays, the mean number of convex hull nodes was 8.35 (SD = 1.41), the mean percentage of internal node intersections was 23.3%, and the mean standard deviation in the internode distance was 0.245.

With these statistics as a basis, we generated a number of random 25-node arrays. Of these, we selected a number that had a Low (13), Medium (16 or 17), or a High (20) number of internal nodes. This selection was also constrained to have a constant proportion (23.0% to 23.6%) of internal node intersections and a constant

(0.24 to 0.25) standard deviation in the internode distance. A further subset of these arrays was then chosen, in which the number of internal nodes was visually indisputable. In all, 18 different arrays were chosen: six from each of the three internal nodes groups. In addition, one from each group was chosen for practice.

For each of these 18 arrays, a benchmark (putatively optimal) closed TSP solution was calculated, using a simulated annealing heuristic (Reinelt, 1994). Random solutions were then generated for each array that had a Low (0 to 5%), Medium (10 to 15%) or High (20 to 25%) difference from the benchmark solution. (The 0 to 25% range of deviations from the benchmark was chosen because it is representative of the range found in the solutions produced by participants in several of our previous TSP studies.)

For each of these random solutions, we noted the number of nodes that were linked to their nearest neighbour in the solution path. The lowest and the the highest numbers of nodes connected to their nearest neighbour that could be found in solutions corresponding to each of the 9 difference from benchmark and interior node group combinations was 16 and 23, respectively. Accordingly, we determined that the three nearest neighbour linked groups should be Low (15 to 16), Medium (19 to 20) and High (23 to 24).

Thus, random solutions were generated for 6 problem instances in each of the three Internal Nodes (IN), three Differences from Benchmark (DB), and three Nearest Neighbour Linked (NNL) groups. These solutions were selected to have no path crossings (and, as a result, connected the convex hull nodes in sequence). More than one solution was generated for each instance in order that unrealistic tours (ones that incorporated obvious, zig-zag backtracking) could be eliminated. However, no other restriction was placed on the solutions. The number of attempts required to generate a single acceptable solution for a given problem instance in a given group ranged from 100 to 1,000,000.

Stimuli. Participants rated the perceived goodness of 162 configurations, each consisting of complete tours (Hamiltonian circuits) of 25 randomly distributed nodes. The 162 configurations were made up of 6 instances drawn from each of three (High, Medium and Low) IN, DB and NNL groups (i.e., of 27 level × group combinations). The nodes were represented by black circular dots, 1.5 mm in diameter, connected by black links on a light grey background and were displayed within a 14 × 14 cm square, with no border, at the centre of a computer screen. Examples of the stimuli are shown in Figure 1.

Insert Figure 1 about here

Design. Each participant was presented with the 162 configurations in a different order that was random, but subject to the constraint that all 27 combinations of number of internal nodes, level of optimality, and number of nodes linked to their nearest neighbours were presented once, before being repeated.

Procedure. Following a practice run of 12 configurations, in which the upper and lower values of each of the three factors were represented, participants carried out the test run, at their own pace, in one unbroken series. Each instance was initiated by the participant by clicking on a “Proceed to Next Test” button. This was followed by a 1-

second "Get Ready" cue, succeeded by a 1-second attention cue, after which the instance and the response boxes were shown on the screen simultaneously.

Participants were asked to give a "rating of the aesthetic appeal of the figure". This involved clicking on one of six response boxes, located along the bottom of the screen and containing both a number and a description. From left to right, these were: "0: very poor"; "1: poor"; "2: fairly poor"; "3: fairly good"; "4: good"; and "5: very good". Each configuration remained until the rating was registered. The entire series of practice and test judgments required around 20-30 minutes to complete.

RESULTS

Figure 2 summarizes the relationship between the aesthetic appeal ratings, averaged over participants, and the 27 different combinations of nearest neighbors, convex hull points, and difference from optimality theoretical measures. The 3x3 grid of panels corresponds to the three levels of nearest neighbors (in rows) and convex hull points (in columns). Each panel shows as a bold line the mean aesthetic rating averaged across all subjects, with one standard error in each direction, as a function of the level of difference from optimality. Each panel also shows, in grey, the mean values for the other panels, to allow for direct visual comparison.

Insert figure 2 about here

It is clear from Figure 2 that changes in the difference from benchmark affect aesthetic appeal, with those configurations that are near-optimal being judged as more appealing. There is also a suggestion of a consistent, but much less pronounced, pattern in relation to the number of convex hull points, with mean aesthetic appeal being greatest for those configuration with the most points on the convex hull.

We adopted a Bayesian approach to statistical inference (e.g., Jaynes 2003) to examine the relationship between aesthetic appeal and the three theoretical measures. Following Vickers, Bovet, Lee and Hughes (2003) and Vickers, Lee, Dry and Hughes (2003), this was done by considering the Bayes Factors for eight linear regression models. These models range in complexity from one where aesthetic judgment is assumed to be constant, and not affected by the theoretical measures, up to a four-term linear combination that includes all three theoretical measures plus a constant. The remaining six models include a subset of one or two of the theoretical measures plus a constant, and have intermediate complexity.

In fitting these models we assumed that the aesthetic appeal judgments follow a Gaussian distribution with common variance, and so used the likelihood function,

$$p(D \mid M, \theta) = \prod_i \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(d_i - \hat{d}_i\right)}{2\sigma^2}\right)$$

where $\hat{d}_i$ is the prediction model by the model M at parameterization θ for the i-th datum. This assumption has the advantage of being sensitive to the precision,

σ, of the empirical data, as measured by the variance across subjects in forming the mean.

We found maximum likelihood fits (corresponding to weighted least-squares optimization of the log-likelihood) for each of the eight models, and evaluated which of the models provided the best account. Because the models have different complexities, deciding which is the best model cannot be done solely through comparing their levels of fit or correlation (Roberts & Pashler, 2000). Accordingly, we used the Bayesian approach to model selection (e.g., Kass & Raftery, 1995; Myung & Pitt, 1997; Pitt, Myung, & Zhang, 2002), which naturally balances the competing demands of fit and complexity.

Computationally, we accomplished this using an index called the Bayesian Information Criterion (BIC; Schwarz, 1978) that is simple to calculate given the fit and parametric complexities of our eight models. For a model, M, with K parameters, and having maximum likelihood parameterization θ^* under the likelihood function $p(\cdot)$ to data D with N samples, the BIC is given by

$$BIC = -2 \log p\left(D \mid M, \theta^*\right) + K \log N.$$

Because the BIC penalizes models for using additional parameters, and is sensitive to data precision, it chooses between models in a way that guards against over-fitting. This means that the theoretical measures it includes in the best linear models can be interpreted as providing some significant capability for explaining and predicting empirical performance. BIC values are also easy to interpret through their relationship to the odds ratios known as Bayes Factors (Kass & Raftery, 1995), which measure how much more likely one model is than another on the basis of the data.

Insert Table 1 about here

Table 1 summarizes the results of this Bayesian analysis, showing the log-likelihood data fit, parametric complexity, BIC value, and Bayes Factor for all eight linear models. The most likely model is the one that includes the difference from benchmark and convex hull theoretical measures. The next most likely models are those that use just the difference from benchmark measure, all of the measures, or the difference from benchmark and nearest neighbor measures. They are respectively about two, five, and eleven times less likely. All of the remaining models are very unlikely, with Bayes Factors well over 100.

Insert Figure 3 about here

Figure 3 shows the relationship between the empirical mean aesthetic judgments and the modeled values provided by the most likely model using the difference from benchmark and convex hull theoretical measures. The correlation between the modeled and empirical values is 0.94.

DISCUSSION AND CONCLUSIONS

Results from the present study provide clear confirmation of Ormerod and Chronicle's (1999) primary finding that participants' ratings of the goodness of TSP solutions were correlated with their relative optimality. However, the present results call into question Ormerod and Chronicle's further conclusion that perceived goodness was also related to figural 'complexity', as measured by the number of interior nodes in each TSP instance. When such factors as symmetry, variability in edge length and the proportion of potential intersections are held constant, then the perceived goodness of TSP solutions is independent of the number of interior nodes, considered as a single factor. It is only if combinations of factors are considered that the results are somewhat better predicted by taking into account the number of interior nodes as well as differences from the benchmark.

Thus, the present results also fail to support the view, suggested by the earlier paper of MacGregor and Ormerod (1996), that the identification of the convex hull is important in determining an overall Gestalt and that, the stronger the tendency to perceive such a global organisation, the more attractive the solution will be judged to be. Taken in conjunction with the finding by Vickers, Lee, Dry and Hughes (2003) that the greater the number of points on the convex hull of randomly generated arrays, the *poorer* were participants' solutions, the present results argue against the influence of the convex hull as an important determinant of both the optimality of participants' solutions and the perceived aesthetic quality of such solutions.

It is not clear what underlies the weak additional contribution by the number of internal nodes, when considered in conjunction with differences from the benchmark. It is possible that aesthetic perception is influenced to some extent by the increased convexity of figures with more nodes on the convex hull. Equally, it may be that such figures are perceived as being slightly more symmetrical. The important point is that the extent of the additional contribution made by this factor is very small.

At the same time, although the present results do not contradict Vickers, Lee et al.'s (2003) proposal that participants find TSP solutions by means of a hierarchical process of linking clusters of nearest neighbours, they do run counter to the hypothesis that it is not the optimality of solutions that is attractive but the extent to which solutions incorporate naturally perceived structure based on nearest neighbours. In other words, although it still seems likely that participants arrive at TSP solutions by exploiting nearest neighbour structures, the attractiveness of such solutions is based on the extent to which they achieve the goal of minimising overall path length, rather than on the means by which they achieve this goal. Although the process of arriving at minimal structures seems to be best understood in terms of a bottom-up, locally based process, the aesthetic satisfaction derived from such structures appears to be based on an overall, global assessment of their economy. That is, the aesthetic appeal of a structure seems to based on an evaluation of the structure as an outcome or solution arrived at by some optimisation process.

This conclusion agrees very closely with some recent - independently derived – psychological and philosophical analyses of beauty (e.g., McMahon, 2000; Ramachandran & Hirstein, 1999). For example, McMahon (2000) has argued that the perceptual principles, identified by Ramachandran and Hirstein (1999) as being involved in the perception of beauty, are analogous to creative problem-solving heuristics. That is, such principles are applicable to problems (such as combinatorial optimisation problems) that are not soluble by means of algorithms or procedures applied in a mechanical way. Similarly, McMahon (2000) has characterised the experience of beauty as "a feeling of clarity … compatible with … experiencing perception as the solution to a problem" (p. 29).

On the other hand, the present results do not provide support for McMahon's further suggestion that the experience of aesthetic attractiveness might be determined by the extent to which a particular structure epitomises the normal constructive processes of perception. In the present experiment, aesthetic judgment does not appear to depend on the extent to which a TSP 'solution' incorporates elements (specifically, the number of nodes on the convex hull) that have been suggested as facilitating global perceptual organisation. Similarly, ratings of attractiveness do not appear to be determined by the extent to which presented structures include nearest neighbour links that have been shown to be important for local organising processes.

This last conclusion may seem counterintuitive. However, it can perhaps be understood from a wider biological perspective. Because the processes of perception do not infallibly give us a veridical picture of our environment, it may be maladaptive for aesthetic pleasure to depend upon a self-confirming correspondence with the operation of these processes. On the other hand, natural structures tend to have evolved as optimal solutions to the problem of survival under the constraints that apply to particular organisms (Hildebrandt & Tromba, 1996; Mandelbrot, 1983). It is therefore not unreasonable to expect the human visual system to have evolved to detect minimal structures. On this view, the aesthetic satisfaction provided by a structure may operate as a metacognitive experience that guides the perceptual process in its locally based, bottom-up process of finding parsimonious structure despite the combinatorial complexity of such a task and the computational challenge of finding an optimal solution. In this respect, the experience of aesthetic attraction seems to resemble the subjective feelings of confidence that accompany many perceptual and cognitive judgments. If this comparison is valid, then it suggests that aesthetic experiences may provide us with a way of evaluating the results of more automatic, low-level brain processes, thereby helping to ensure that we function in an adaptive way that is well calibrated with respect to the most parsimonious and most likely description of some relevant aspect of the world. It is perhaps this adaptive function that underlies the importance assigned to aesthetic considerations by many mathematicans and scientists in solving problems in a new way (McAllister, 1996).

NOTES

1. Potential intersections can be counted for all possible pairs of links between all possible pairs of nodes (without any pair of links having a common node). The proportion of pairs of links that result in intersections and involve all possible nodes remains roughly constant (at around 23.2% of all possible pairs of links), irrespective of the number of nodes in an array. Similarly, the number of pairs of links that result in intersections and that involve internal nodes only remains a roughly constant proportion (around 23.3% of all possible links between internal nodes). This proportion remains roughly constant, even when the number of nodes on the convex hull varies. In contrast, the number (and proportion) of intersections involving all possible nodes increases as a function of the number of nodes on the convex hull. For this reason, it was decided to manipulate the number of potential intersections involving internal nodes only, so as to minimise any possible interaction between the proportion of intersections involving internal nodes and the number of nodes on the convex hull.

REFERENCES

Ahuja, R.K., Magnanti, T.L., & Orlin, J.B. (1993). *Network flows: Theory, algorithms, and applications.* New York: Prentice Hall.

Attneave, F. (1982). Prägnanz and soap bubble systems: A theoretical exploration. In: J. Beck (Ed.), *Organization and representation in perception*. Hillsdale, NJ: Lawrence Erlbaum, pp. 11-29.

Caelli, T.M. (1981). Some psychophysical determinants of discrete Moiré patterns. *Biological Cybernetics*, *39*, 97-103.

Dawson, M.R.W. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review*, *98*, 569-603.

Hildebrandt, S. & Tromba, A. (1996). *The parsimonious universe: Shape and form in the natural world*. New York: Springer-Verlag.

Hwang, F.K., Richards, D.S., & P. Winter, P. (1992). *The Steiner tree problem. Volume 53 of Annals of Discrete Mathematics.* Amsterdam, Netherlands: North-Holland.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. New York: Cambridge University Press.

Kass, R.E., & Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773-795.

MacGregor, J.N. & Ormerod, T.C. (1996). Human performance on the traveling salesman problem. *Perception and Psychophysics*, *58*, 527-539.

McAllister, J. (1996). *Beauty and revolution in science*. Ithaca, NY: Cornell University Press.

McMahon, J.A. (2000). Perceptual principles as the basis for genuine judgments of beauty. *Journal of Consciousness* Studies, *7*, 29-35.

Mandelbrot, B.B. (1983). *The fractal geometry of na*ture. New York: W.H. Freeman & Co.

Myung, I.J., & Pitt, M.A. (1997). Applying Occam's razor in modeling cognition. *Psychonomic Bulletin & Review*, *4*, 79-95.

Ormerod T.C., & Chronicle, E.P. (1999). Global perceptual processing in problem solving: the case of the traveling salesperson. *Perception and Psychophysics*, *61*, 1227-1238.

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review 109*(3), 472-491.

Reinelt, G. (1994). The traveling salesman: Computational solutions for TSP applications. *Lecture Notes in Computer Science*, *840*. Berlin: Springer-Verlag.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*, 358-367.

Stein, D.L. (1989). Spin glasses. *Scientific American*, *261*, 36-42.

Vickers, D., Bovet, P., Lee, M.D., & Hughes, P. (2003). The perception of minimal structures: Performance on open and closed versions of visually presented Euclidean Travelling Salesperson problems. *Perception*, *32*, 871-886.

Vickers D., Butavicius, M., Lee, M.D., & Medvedev, A. (2001). Human performance on visually presented traveling salesman problems. *Psychological Research*, *65*, 34-45.

Vickers, D., Dry, M., Lee, M.D., & Hughes, P. (submitted). The role of nearest neighbours in the perception of Glass patterns. *Perception*.

Vickers, D., Lee, M.D., Dry, M., & Hughes, P. (2003). The roles of the convex hull and the number of potential crossings upon performance on visually presented traveling salesperson problems. *Memory and Cognition*, *31*, 1094-1104.

Vickers, D., Mayo, T., Heitmann, M., Lee, M.D., & Hughes, P. (in press). Intelligence and individual differences in performance on three types of visually presented optimisation problems. *Personality and Individual Differences*.

Vickers, D., Preiss, A.K., & Hughes, P. (submitted). The role of nearest neighbours in the perception of structure and motion in dot patterns. *Perception*. (Under review).

Zahn, C.T. (1971). Graph-theoretical methods for detecting and describing Gestalt clusters. *IEEE Transactions on Computers*, *C-20*, 68-86.

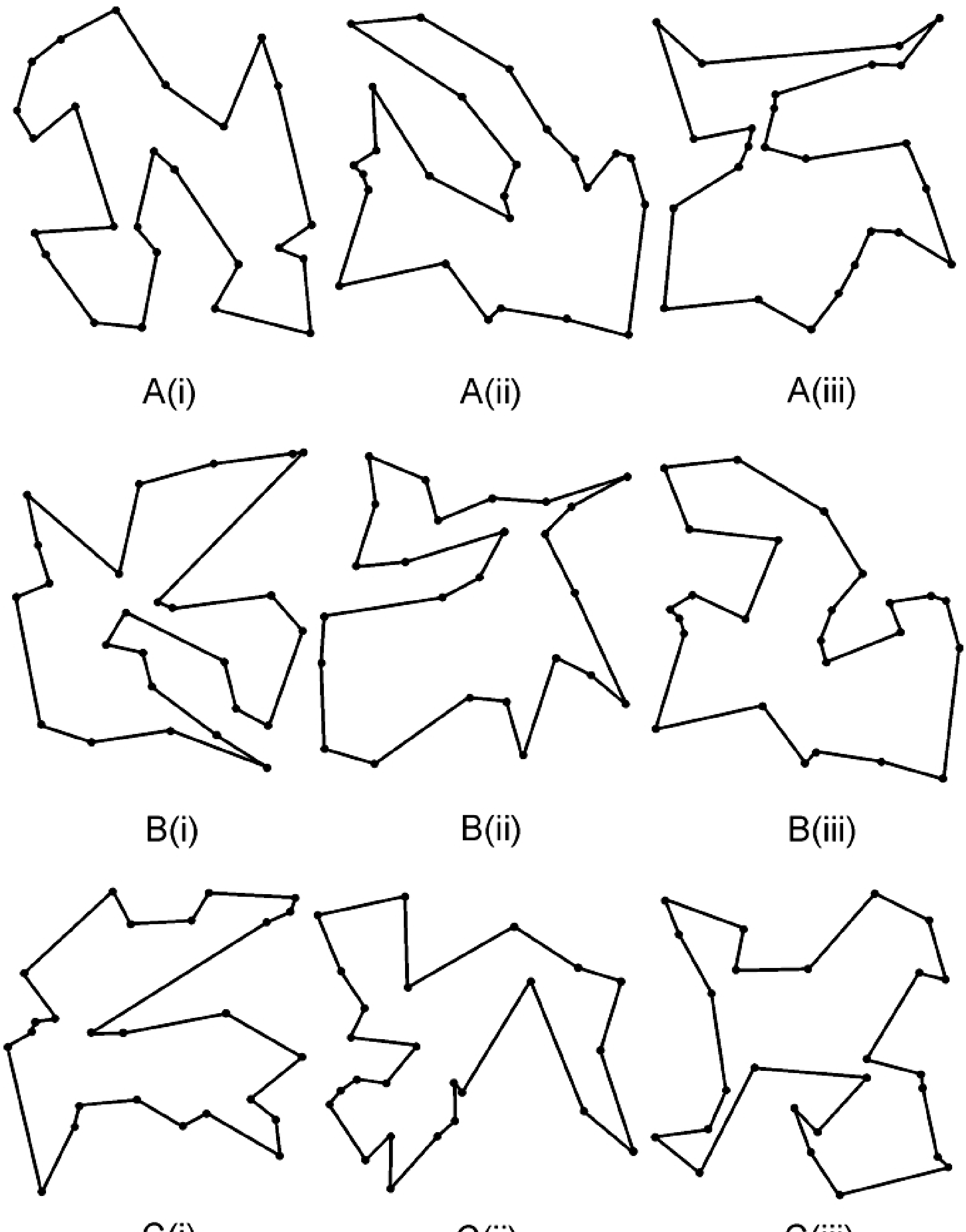
FIGURES

Figure 1. Examples of the stimuli presented to participants. Figures 1A(i) to 1A(iii) all show examples of stimuli with a Medium deviation from the benchmark and a Medium number of nodes linked to their nearest neighbours. These Figures show instances of stimuli with a Low, a Medium and a High number of interior nodes, respectively. Figures 1B(i) to 1B(iii) all show examples of stimuli with a Medium number of interior nodes and a Medium number of nodes linked to their nearest neighbours. These Figures show instances of stimuli with a Low, a Medium and a High Medium deviation from the benchmark, respectively. Figures 1C(i) to 1C(iii) all show examples of stimuli with a with a Medium number of interior nodes and a Medium deviation from the benchmark. These Figures show instances of stimuli with a Low, a Medium and a High number of number of nodes linked to their nearest neighbours, respectively.

Figure 2.

Figure 3.

Figure 1, Vickers, Lee, Hughes, Dry & McMahon



Model	Log-Likelihood Data Fit	Parametric Complexity	BIC Value	Bayes Factor
Constant	-48.26	**1**	99.81	>100
DO	21.59	**2**	-36.59	2.42
CH	-48.25	**2**	103.09	>100
NN	-48.25	**2**	103.10	>100
DO+CH	24.12	**3**	-38.36	1
DO+NN	21.66	**3**	-33.44	11.71
CH+NN	-48.25	**3**	106.38	>100
DO+CH+NN	24.16	**4**	-35.14	5.01

Table 1. The log-likelihood data fit, parametric complexity, BIC value, and Bayes Factor for all possible linear models of mean aesthetic appeal judgments using the theoretical measures (DO=deviation from optimality, CH=convex hull points, NN=nearest neighbors). The Bayes Factors are taken in relation to the most likely model, which is the DO+CH model.