Psychological Research (2001) 65: 34–45

ORIGINAL ARTICLE

Douglas Vickers · Marcus Butavicius · Michael Lee
Andrei Medvedev

Human performance on visually presented Traveling Salesman problems

Received: 28 December 1998 / Accepted: 20 January 2000

Abstract Little research has been carried out on human performance in optimization problems, such as the Traveling Salesman problem (TSP). Studies by Polivanova (1974, *Voprosy Psikhologii*, *4*, 41–51) and by MacGregor and Ormerod (1996, *Perception & Psychophysics*, *58*, 527–539) suggest that: (1) the complexity of solutions to visually presented TSPs depends on the number of points on the convex hull; and (2) the perception of optimal structure is an innate tendency of the visual system, not subject to individual differences. Results are reported from two experiments. In the first, measures of the total length and completion speed of pathways, and a measure of path uncertainty were compared with optimal solutions produced by an elastic net algorithm and by several heuristic methods. Performance was also compared under instructions to draw the shortest or the most attractive pathway. In the second, various measures of performance were compared with scores on Raven's advanced progressive matrices (APM). The number of points on the convex hull did not determine the relative optimality of solutions, although both this factor and the total number of points influenced solution speed and path uncertainty. Subjects' solutions showed appreciable individual differences, which had a strong correlation with APM scores. The relation between perceptual organization and the process of solving visually presented TSPs is briefly discussed, as is the potential of optimization for providing a conceptual framework for the study of intelligence.

Introduction

The Traveling Salesman problem (TSP) is a well-known type of combinatorial optimization problem, belonging to the class of so-called NP-complete problems, for which it is believed there is no algorithm that can be guaranteed to arrive at an optimal solution within a practicable (polynomial) time (Lawler, Lenstra, Rinooy Kan, & Shmoys, 1985). That is, no algorithm can solve the problem in a time proportional to n^c (or better), where n is the number of relevant input variables and c is some constant (Goldschlager & Lister, 1988). A common version of the TSP is formulated as follows: Given a set of n cities, and a specified cost (or distance) incurred in traveling between any two of them, devise an itinerary, such that (a) each city is visited exactly once, and (b) the total cost (or distance) is kept to a minimum. To solve this problem definitively involves considering $(n\text{-}1)!/2$ pathways. For 5 cities, this amounts to 12 distinct pathways. With 10 cities, the 181, 440 possibilities are still manageable. However, for just 25 cities, the number of pathways is so immense that a computer evaluating 1 million possibilities a second would take 9.8 billion years (or two thirds the age of the universe) to evaluate them all (Stein, 1989).

Such optimization problems arise in many different, often unexpected contexts. In our case, we were looking for a simple neuropsychological test that could be administered repeatedly to monitor recovery during hyperbaric treatment. One favored candidate was the Trail Making Test (TMT) from the Halstead-Reitan Neuropsychological Battery (Reitan, 1992). There are two forms in the TMT. In form A, there are 25 circles, numbered from 1 through 25, 'randomly distributed' on a page. Subjects have to draw a path, connecting them up, as quickly as possible, in the order 1, 2, 3, … . In form B, there are numbers and letters, and the subject has to connect the 25 circles in the order 1, A, 2, B, 3, C, … . In both cases, the score is the time taken to complete the path. The TMT correlates around 0.5 with psychometric measures of intelligence, is widely regarded as a sensitive indicator of brain damage (Lezak, 1995), and has a claim to being the single, most frequently used neuropsychological test (Butler, Retzlaff, & Vanderploeg, 1991).

D. Vickers (✉) · M. Butavicius · M. Lee · A. Medvedev
Psychology Department, University of Adelaide,
Adelaide 5005, South Australia

Unfortunately, a major disadvantage with the TMT is that there is no algorithm for generating theoretically (or measurably) equivalent, but stochastically different, forms for repeated testing. Noting that the pathways employed in both forms of the TMT were not random, but took the form of constrained, self-avoiding curves, we compared their lengths with those of 10,000 random pathways through the same sets of points. We found that the TMT pathways were between 3 and 5 standard deviations shorter than the means of the distributions of random itineraries. We, therefore, suggested that the pathways in forms A and B constituted near-optimal routes through the 25 circles in each case. On this basis, we proposed a method for generating multiple test forms, using an elastic net algorithm (Durbin & Willshaw, 1987; Peterson & Söderberg, 1995) to produce near-optimal pathways linking sets of 25 randomly distributed circles (Vickers & Lee, 1998). A similar approach, using a Hopfield-style neural network, also proved useful in generating pathways for an alternative path-following test – the so-called 'number-joining' or Zahlen-Verbindungs-Test (Oswald & Roth 1987), for which the inter-nodal distances could not be represented by a metric arrangement of the nodes in question (Lee, Brown, & Vickers, 1997).

Human performance on TSPs

The possibility that performance on the TMT might depend – at least partly – on how easily subjects detect optimal pathways raises the further question of how well human beings might solve visually presented optimization problems like the TSP. In addition, there are good reasons for investigating performance on optimization problems. For instance, if a major function of the brain is to provide the individual with a model of his or her environment, then we should expect the most general physical principles operating in the natural world would be reflected in the way the brain constructs such a model (Shepard, 1984).

One such constraint is the 'minimum principle'. This states that, in passing from one state to another, physical systems tend to expend the minimum amount of energy. Some form of minimum principle has been invoked to account for a diverse array of phenomena. These include the shapes of soap bubbles, the structure of spider webs, and the circular plan of walled cities (Hildebrandt & Tromba, 1996). Many architectural constructions, such as Frei Otto's Olympic Stadium in Munich, are intended to achieve strength with a minimum of material. Similarly, as exemplified in the foraging patterns of animals, survival in the natural world frequently depends upon the capacity of an organism to minimize the value of some function, such as energy expenditure, under a multiplicity of constraints (Stephens & Krebs, 1986).

Optimization problems, such as foraging, elude exact solution by means of deterministic algorithms, and a separate mathematical discipline of 'optimization' (and a corresponding branch of computer science) has developed to find solutions to such challenges. At the same time, the ability to arrive quickly at near-optimal solutions to such problems seems to be a characteristic of much human perception and cognition, and is often taken to epitomize ingenuity and intelligence, as opposed to sheer mechanistic computation. When a chess master defeats a powerful computer, this is hailed as a triumph of human intelligence over brute-force serial computation. In contrast, the same intelligence is not ascribed to the calculating feats of people historically labeled as 'idiots savants' (Anderson, 1992).

In psychology, there have been several attempts to formulate some kind of minimum principle for human behavior and mental activity, a well-known example being Zipf's (1949) book, *Human behavior and the law of least effort*. Echoing this, the 'Prägnanz principle', enunciated by the German Gestalt school, proposed that we tend to perceive that structure which is minimal, in the sense of being simplest and most economical (Köhler, 1929). Some form of optimization hypothesis has long underpinned the notion of rational behavior in economics and decision making. More recently, Anderson (1990, 1991) has argued that human cognition in general can be seen as an optimal response to the information-processing demands of the environment.

Given the prevalence of optimizing processes, it might be expected that our capacity to perceive minimal structures and to solve optimization problems would constitute a major focus of research. However, our first literature search found only one paper (in Russian) that focussed on the optimization process itself. In an exploratory study, Polivanova (1974) presented subjects with TSPs in two formats. In the first, she gave the subject a table of all the distances between (say) 10 randomly distributed points. In the second, she gave the subject a diagram with the points marked on it. In both cases, the subject had to find a path that would go through each point once only and return to the starting point.

Polivanova employed simple problems, with 4, 6, or 10 points only, and, as might be expected, her subjects did much better with the visual format. Polivanova speculated that this was because it was easier to apply heuristics to the visual problems. For example, subjects said they looked for 'simple', 'convex' forms, and avoided pathways that crossed. They also looked for pathways that seemed 'natural' or 'aesthetically pleasing'.

As we were preparing to investigate these possibilities, a second article on the same topic appeared. MacGregor and Ormerod (1996) looked at performance on 10- and 20-point problems – all presented as visual arrays. They found that their subjects' solutions were all close to the best known solutions. In consequence, they found no individual differences and a zero correlation between performance across different arrays. They concluded that detecting minimum paths was an innate, natural tendency of the visual system, like seeing the world in three dimensions.

MacGregor and Ormerod (1996) also suggested that subjects make use of the principle, enunciated by Flood (1956), that the optimal solution will always connect adjacent points on the convex hull in sequence (even though it also passes through interior points in the process). The convex hull is a boundary, such that no line joining any two points in the array can fall outside it. Specifically, MacGregor and Ormerod argued, the difficulty of finding a solution will depend, not on the *total* number of points, but on the number of points *on* the convex hull (or, equivalently, on the number of points *inside* the convex hull).

This last conclusion seems counterintuitive. In addition, Lee and Vickers (2000) have questioned it on three counts: First, the nodal points in the arrays of MacGregor and Ormerod (1996) were not randomly distributed. To manipulate the number of points on the convex hull, MacGregor and Ormerod began with a regular polygon with the requisite number of points on the hull. They then subjected the vertices to a random angular displacement of between plus and minus 5 degrees. This means that, with the 10-point arrays, it would be possible to draw a circle that would pass through between 4 and 9 boundary points. In the case of the 20-point arrays, the circle would pass through between 4 and as many as 16 boundary points. It is difficult to believe that subjects would not pick up the high degree of rotational symmetry in several of these arrays (Glass, 1969; Pickover, 1984). Second, after fixing the boundary points, MacGregor and Ormerod generated the interior points randomly, with the constraint that they fell within either a circular or a torus-shaped area inside the hull. As shown by Lee and Vickers, the probability that a random configuration of nonboundary points will fall entirely within the prescribed areas inside the convex hull is negligibly small in most cases. Again, it seems likely that the subject would be sensitive to these constraints, and that this would favor a convex hull strategy. Third, the influence of the number of boundary points is, in principle, restricted. To demonstrate this, Lee and Vickers generated 10,000 simulated arrays with 10, 20, 30 … 100 total points. According to these simulations, the number of boundary points increases very little beyond 50 points. For example, the 20-point arrays, used by MacGregor and Ormerod, would be expected to have only 7 or 8 points on the convex hull; any array with 10 or more points (or 6 or fewer) would be extremely unusual.

These constraints resulted in arrays that were not only non-random but may well have increased the saliency of the convex hull for the subjects, thereby biasing the findings in line with the experimental hypothesis of MacGregor and Ormerod (1996).

Experiment 1

The first experiment was carried out to examine the effect of the number of points on the convex hull on the optimality of solutions to visually presented TSP arrays. A second aim was to explore the possibility, suggested by both Polivanova's subjects and by MacGregor and Ormerod, that the perception of optimal structure might be a natural, automatic tendency of the human visual system, as opposed to a specific, task-determined and capacity-limited achievement.

Method

Subjects. The 36 subjects came from a variety of backgrounds and educational levels.

Stimuli. In each stimulus array, the points to be linked were uniformly distributed within a rectangular area. There were six arrays in all. Two had 10 points, two had 25, and two had 40 points. Each pair of arrays had either a high or a low number of points falling on the convex hull.

The number of points on the convex hull was determined by first generating 10,000 random arrays with a given number of points, and calculating the mean number of points that fell on the convex hull for each array size. We then chose an array with a number that was 1 SD above this, and one that was 1 SD below. This meant that the two 10-point arrays had either 5 or 7, the two 25-point arrays had 7 or 9, and the two 40-point arrays had either 8 or 12 points on the convex hull.

Procedure. Subjects were partitioned randomly into two equal-sized groups, labeled O and G. Both groups were then presented with the six stimulus arrays described above.

For each array, the 'Optimization' group (O) were instructed to draw a pathway through all of the points, such that: (a) the path passed through each point once and once only; (b) the overall pathway was as short as possible; and (c) the completed pathway returned to the starting point. This constitutes the classical TSP.

The corresponding instructions for the 'Gestalt' group (G) were to draw a pathway through all of the points, such that: (a) the path passed through each point once and once only; (b) the overall pathway looked most natural, attractive, or aesthetically pleasing; and (c) the completed pathway returned to the starting point. Having completed the six arrays, subjects in both groups were asked to rate the difficulty of the task, and to describe briefly how they set about the task. Difficulty ratings were made by marking a linear scale labeled "−5" (extremely difficult), through "0" (intermediate), up to "5" (extremely easy).

Subjects in both groups were allowed to start their pathways at any point. They were permitted to erase and redo any part of any pathway, but were asked to try to avoid this if possible. The time taken to complete each array was recorded, but subjects were assured that this was not a test of speed, and that they should take as much or as little time as they wanted.

Results

A connectionist model for human performance on the TSP

Comparative solutions to the TSP problems were found using two main approaches. The first employed the 'elastic net' approach of Durbin and Willshaw (1987) as a potential candidate for the human solution process. The net is made up of a number of elements, chosen to be two or three times greater than the number of cities (Durbin, Szelski, & Yuille 1989, Simmen 1991). Each element is continually acted upon by two forces. The

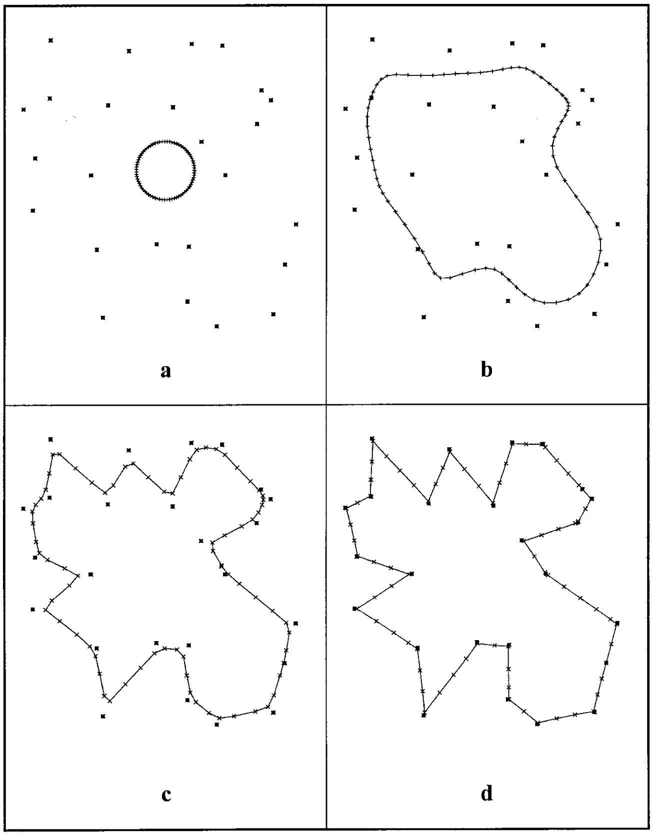

Fig. 1 Four successive stages in the fitting process of the elastic net algorithm, from start (**a**) to finished solution (**d**)

first force acts to keep each net element near its topological neighbors. The second force, which is given a greater weighting, attracts elements to position themselves on cities, with an even greater force of attraction being applied for those elements closest to a given city. Durbin and Willshaw (1987) note that the combined action of these forces performs a type of gradient-descent optimization, with the result that the final configuration of the net corresponds to a (local) minimum of a global error function. Figure 1 illustrates four stages in the progressive effect of these forces for a set of 25 randomly located cities. As with other applications of the elastic net reported here, this example used similar parameter values to those suggested by Durbin and Willshaw (1987), and experimentation with other parameter regimes did not produce better results.

Although hypothetical, this parallel, analogue algorithm accords well with current connectionist approaches in cognitive research, which focus on the global activity of a large number of 'neuron-like' units. Indeed, it is derived from a biological mechanism postulated to explain the development of neuronal structures associated with vision in the brain (Malsburg & Willshaw, 1977; Willshaw & Malsburg, 1979). In addition, the 'elastic net' algorithm represents a holistic, geometrical approach to the TSP, which concurs with the 'gestalt-like', visual-based problem solving style suggested by previous research. It is simple, robust, and versatile (Yuille, 1995), and its performance, both in terms of speed and proximity to optimal solutions, is comparable to more conventional and serial TSP algorithms (Durbin & Willshaw, 1987) and alternative parallel distributed approaches (Peterson, 1990).

Calculation of optimal or near-optimal pathway lengths

In the second approach, we followed MacGregor and Ormerod (1996) in their use of the TRAVEL suite of algorithms, designed to produce provably good solutions to TSPs (Boyd, Pulleyblank, & Cornuéjols, 1987). This comprises six conventional, serial TSP algorithms, including the Nearest Neighbor, Largest Interior Angle, and Convex Hull heuristics (Norback & Love, 1977; Golden, Bodin, Doyle and Stewart, 1980), as well as several well-proven tour improvement routines, including the Lin and Kernighan (1973) heuristic. The optimality of subjects' pathways was determined by comparing them with lower bounds estimated by the TRAVEL program by means of a number of different methods. For example, one such estimate can be obtained by calculating the distance between each point in the array and its nearest neighbor, and summing the distances. This may not correspond to a valid pathway. However, no valid pathway can be shorter than this lower bound. Consequently, if a pathway is found, which is as short as this, we know that this is *an* optimal solution (although not necessarily the *only* optimal solution).

All of the above algorithms were employed in the search for optimal solutions for each of the six arrays, although no single algorithm consistently provided the shortest pathway. By these means, we were able to arrive at optimal solutions for five of the stimulus arrays. In the case of the sixth (the 40-point problem with 12 points on the convex hull), the best solution was conservatively estimated to be no more than 0.2% from the lower bound, so that this solution may or may not be optimal. In this case, the best solution was adopted as a surrogate for the optimal.

Having estimates of optimal pathway lengths meant that we could express the subjects' pathway lengths for each problem in a standard, normalized form, as done by MacGregor and Ormerod (1996). These authors expressed each pathway length in terms of the standard deviation of a distribution of 100 random pathways through the same set of points. Because our stimuli differed in number of points, we also wished to standardize over problems of differing size and difficulty. This was accomplished by dividing the *difference* between the length of the optimal pathway for each array and that of the subject's pathway by the standard deviation in length for 10,000 random pathways through the same number of points. As shown by Vickers and Lee (1998), these distributions are highly symmetric, and closely resemble normal distributions. (When distributions of 10,000 random pathways through each of the six problem arrays were examined, the average measure of skewness was −0.12 and the average measure of kurtosis was 2.88. These figures are sufficiently close to the respective values of 0 and 3, for a perfectly normal distribution, to justify the assumption of normality.) The resulting measure, z_{opt}, allows a comparison across problems differing in number of points and geometrical complexity, and takes into account the fact that, at each problem size, subjects' solutions were tightly bunched towards the optimal value. Thus, an optimal solution would have a score of $z_{opt} = 0$. Any other score is necessarily positive, and represents a standardized measure of the difference between the pathway length and the shortest possible pathway length for that problem array.

For similar reasons, the times taken to complete the arrays were converted to a speed score, *s* (the inverse of the time taken per point).

Speed of completing stimulus arrays

A repeated measures ANOVA was carried out, with *s* as the dependent variable, a between-subjects factor of Group (G or O), and two within-subjects factors: Convexity (a HIGH or LOW number of points on the convex hull); and Number (the total number of points in the array, being either 10, 25, or 40).

The trend analysis of Number showed a highly significant linear component [$F(1, 34) = 33.14$; $p < 0.001$], with a less significant nonlinear component [$F(1, 34) = 9.02$; $p < 0.05$]. The trend was monotonic, but the rate of increase in speed appeared to fall off with increasing problem size ($M_{10\ point\ problems} = 0.53$, SD = 0.31; $M_{25\ point\ problems} = 0.80$, SD = 0.33; $M_{40\ point\ problems} = 0.90$, SD = 0.43).

There was a significant effect of Convexity [$F(1, 34) = 11.44$; $p < 0.05$, $M_{high\ convexity\ problems} = 0.78$, SD = 0.31 *vs* $M_{low\ convexity\ problems} = 0.70$, SD = 0.33], with subjects in both groups completing arrays faster when there was a greater number of points on the convex hull [except for group O, who performed more slowly with the 40-point, HIGH array (M = 0.71, SD = 0.46) than with the corresponding LOW array (M = 0.78, SD = 0.44)].

There was a significant effect of Group [$F(1, 34) = 5.00$; $p < 0.05$], with the G group (M = 0.85, SD = 0.31) being somewhat faster overall than the O group (M = 0.63, SD = 0.29).

There were also two significant interactions: a two-way interaction between Convexity and Number [$F(1, 34) = 8.63$; $p < 0.05$]; and a three-way interaction between Group, Convexity, and Number [$F(1, 34) = 8.35$; $p < 0.05$]. The first indicates a relatively greater increase in speed from LOW to HIGH convexity as the number of points is increased ($M_{10\ low} = 0.51$, SD = 0.28 *vs* $M_{10\ high} = 0.54$, SD = 0.37; $M_{25\ low} = 0.70$, SD = 0.38 *vs* $M_{25\ high} = 0.90$, SD = 0.33; $M_{40\ low} = 0.89$, SD = 0.48 *vs* $M_{40\ high} = 0.91$, SD = 0.46). The second is largely a reflection of the anomalous performance of O group with the 40-point, HIGH Convexity array.

Pathway lengths

A similar repeated measures ANOVA was carried out, but with z_{opt}, rather than *s*, as the dependent variable.

Not surprisingly, each Convexity and Number combination had very different variances in the two groups ($p < 0.001$, in all cases), with the G group showing the greater variance in all cases.

There was a significant Group difference in the mean pathway length [$F(1, 34) = 11.28$; $p < 0.05$], with the O group producing shorter pathways on average ($M_{optimization} = 0.30$, SD = 0.20 *vs* $M_{gestalt} = 0.20$, SD = 1.1). Despite the non-normality of the distributions and the apparent violation of the homogeneity of variance assumption, this result is unlikely to have occurred by chance.

There was also a marginally significant non-linear component in the trend of Number [$F(1, 34) = 4.26$; $p < 0.05$, $M_{10\ point\ problems} = 0.81$, SD = 1.16 *vs* $M_{25\ point\ problems} = 0.59$, SD = 0.77 *vs* $M_{40\ point\ problems} = 0.77$, SD = 0.97]. However, this result may well be spurious, given the marginal significance and the violation of assumptions noted above.

No other main or interaction trends or effects were significant, even at the 0.05 level.

Correlations between performance on different stimulus arrays

Matrices of Spearman rank-order correlations between pathway lengths for each array were calculated for groups G and O separately. In the case of group G, the rank order correlations between all 15 possible pairs of arrays were all positive and significant ($p \geq 0.43$; $p \leq 0.039$), with an average *p* of 0.81, and with Kendall's coefficient of concordance, $w = 0.84$ ($df = 5$; $p \ll 0.001$). For group O, the rank order correlations between all fifteen possible pairs of arrays were all positive, with six being significant ($p \geq 0.40$; $p \leq 0.05$), an average *p* of 0.87, and with Kendall's coefficient of concordance, $w = 0.89$ ($df = 5$; $p \ll 0.001$). These results are consistent with the existence of reliable individual differences between subjects.

Pathways common to different subjects and to both groups G and O

Table 1 shows the degree to which certain solutions were repeated by subjects. 'Rank' refers to the rank order in the length of a given pathway, whether produced by a subject in either group G or group O, with 1 denoting the shortest pathway produced by any subject. 'Number of times' refers to the frequency with which a particular solution was produced by different subjects. 'Relative proportion of Gestalt pathways' gives the proportion (in percentage form) of the number of times the pathway in question occurred in the G group, divided by the total number of identical pathways (whether from group G or group O).

Table 1 Relative frequency of common solutions in Experiment 1

Problem	Solution rank	No. of times solution reoccurred	Relative no. of Gestalt solutions (%)
10-7[a]	1	13	30.8
	2	10	40
	3	2	100
10-5[a]	1	8	12.5
	2	6	50
	6	2	0
	16	2	50
25-9[a]	1	13	23
	2	5	40
	8	2	50
25-7a	1	4	0
	2	2	0
40-12[a]	3	2	0
40-8[a]	No common solutions		

[a] Indicates total number of points in the array followed by number of points on the convex hull, e.g., the 10–7 problem has 10 points in total and 7 on the convex hull

The data in Table 1 bring out four main points. Firstly, the lower the number of points, the more overlap there was between the pathways. For example, on the 10-7 array (10 points, with 7 on the convex hull), 25 subjects produced one of the three shortest pathways, while, at the 40-8 level, no two subjects produced the same pathway. Secondly, the closer a particular pathway is to the optimal, the higher the frequency with which that pathway is produced. In other words, if two different subjects produce the same pathway, the chances are that this pathway is going to be a relatively short one. Thirdly, the probabilities that the same pathway would have been reproduced by the observed number of different subjects, if the pathways were being drawn in a random manner, can be regarded as zero for all practical purposes. Fourthly, although most of the repeated pathways were found in group O, a sizable proportion (around 30.5% overall) occurred when subjects were asked to draw purely aesthetic pathways. In effect, 9 of the 13 repeated pathways were produced by *both* groups. That is, several subjects, when presented with the same arrays, but with instructions to maximize aesthetic appearance or to minimize pathway length, produced exactly the same pathways out of a total set of potential pathways that, even for a 10-point array, consists of 181,400 possibilities.

Path uncertainty

Following MacGregor and Ormerod (1996), a measure of 'path uncertainty' was calculated for each array. This was done by constructing an $n \times n$ matrix and counting the frequency with which each point in an array was connected by subjects to each of the other points. These frequencies were then divided by the total number of subjects (18, in each case) to provide probabilities, p_i. The total path uncertainty, H, associated with a given array was then calculated by means of the standard information theory formula,

$$H = \sum_{i=1}^{k} p_i(-\log_2 p_i),$$

where k is the total number of connections made by subjects (Shannon & Weaver, 1949). The convention was followed that zero probabilities give rise to zero uncertainty measures (i.e., if something does not happen, it is not regarded as informative).

The measure of path uncertainty provides a concise description of the variety of different solutions generated by all subjects on a particular array. Following Garner's (1970) principle that good patterns have few alternatives, MacGregor and Ormerod suggest that it may provide a useful first approximation to a measure of the difficulty of each problem.

Path uncertainty measures, calculated for the six arrays were 5.59, 10.22, 9.9, 27.99. 48.99, and 55.46, respectively, for arrays with 3, 5, 16, 18, 28, and 32 points interior to the convex hull. The product-moment correlation between path uncertainty and the number of interior points was $r = 0.94$ ($df = 4$), which is significant beyond the 0.01 level (two-tail test). This virtually coincides with a similar correlation of $r = 0.93$, obtained by MacGregor and Ormerod (1996), between these two quantities.

Comparisons between subjects' pathways and elastic net solutions

Table 2 shows the shortest pathways of the subjects in comparison with the solutions produced by the elastic net algorithm for the same stimulus arrays. All pathway lengths are expressed as the difference between the given pathway length and the optimal (or surrogate) length, expressed, in turn, as a percentage of the optimal (or surrogate) length.

From Table 2 it can be seen that, with the 10-point arrays, there is no difference between the best solutions for the elastic net and the shortest pathways produced by subjects, with both arriving at an optimal pathway. However, as the number of points in the array increases, the performance of the 'best' subjects declines at a much

Table 2 Summary of best solutions expressed in terms of percentage deviation from the optimal (*O group* Optimization group, *G group* Gestalt group)

Problem	Human O group	Human G group	Elastic net
10-7	**0**	**0**	**0**
10-5	**0**	**0**	**0**
25-9	**0**	**0**	1.69
25-7	**0**	3.53	4.11
40-12	2.23	**0.08**	8.84
40-8	**2.69**	4.66	8.1

Note: shortest human tours are highlighted in **bold** slower rate than that of the elastic net algorithm. Unlike the best subjects, the elastic net algorithm failed to find an optimal solution to the 25-point problems, and, with 40-point problems, the algorithm is at least 8% poorer than the optimal. This compares unfavorably with the best human performance, which is at most 4.7% poorer.

When we compare the shortest pathways produced by subjects in groups G and O, those produced in accordance with aesthetic preference were longer (around 3%, on average) only in the case of the 25-7 and 40-8 problems. The shortest pathways, produced by the group G subjects, corresponded to the optimal pathways for three of the arrays, and equaled the performance of the best subjects in group O. Indeed, on the 40-12 array, the shortest pathway produced by a subject in group G was actually *shorter* than the best solution produced by any subject in group O.

Ratings of difficulty

Generally, subjects seemed to find the tasks more easy than difficult, with an average rating, over both groups and all subjects, of 1.64 (SD = 2.25). A trend analysis and repeated measures ANOVA was carried out, similar to the two previous analyses, except for the substitution of the difficulty rating, *d*, as the dependent variable. Not surprisingly, the multivariate test for homogeneity of dispersion was significant at the 0.05 level, suggesting a minor violation of this assumption.

The analysis showed a significant linear increase in perceived difficulty as the number of points in the arrays increased [$F(1, 34) = 11.92$; $p < 0.05$, $M_{10\ point\ problems} = 2.63$, SD = 2.13 *vs* $M_{25\ point\ problems} = 1.72$, SD = 2.09 *vs* $M_{40\ point\ problems} = 1.5$, SD = 2.4], with a significant non-linear component [$F(1, 34) = 4.59$, $p < 0.05$], indicating that the rate of increase in difficulty ratings is not steady, but falls off as the number of points in an array increases.

There was also a significant two-way interaction between Number and Convexity, with a marginally significant linear component [$F(1, 34) = 4.91$; $p < 0.05$], and a highly significant non-linear component [$F(1, 34) = 8.77$; $p < 0.05$]. This reflected the fact that perceived difficulty for the two 10-point arrays was virtually identical ($M_{high\ convexity} = 2.64$, SD = 2.17 *vs* $M_{low\ convexity} = 2.63$, SD = 2.29), and, for the 25-point arrays, the array with more points on the convex hull was perceived as easier ($M_{high\ convexity} = 2.13$, SD = 2.17 *vs* $M_{low\ convexity} = 1.32$, SD = 2.34), while, for the 40-point arrays, it was the array with fewer points on the convex hull that was rated as easier ($M_{low\ convexity} = 1.79$, SD = 2.32 *vs* $M_{high\ convexity} = 1.13$, SD = 2.73).

Reported strategies

A total of seven subjects (from both groups G and O) reported using some form of "circular" strategy, whereby they began by forming a roughly circular perimeter around the points. Eight subjects (again from both groups) reported that the pathway was immediately and intuitively obvious, and seemed to be automatically triggered by the points. However, the G group showed a much wider range of approaches. These ranged from pathway minimization, through avoidance of self intersections, and the imposition of repetitive or spiral patterns, up to one or two that seemed designed to maximize the pathway length.

Figure 2, for example, shows two contrasting pathways drawn by subjects from the G group; Fig. 2a, which has a clear spiral structure, is some 70% longer than the optimal pathway length, while Fig. 2b is a mere 0.08% longer than the optimal – and was shorter than any pathway produced by a member of group O. (Interestingly, Fig. 2b was drawn by a fashion designer, who also produced the shortest pathway among the subjects for all other arrays, except the 25-9 configuration, for which she produced the second shortest pathway.)

Figure 2c shows the solution produced by the elastic net algorithm to the 40-12 array, also shown in Fig. 2b. This solution was over 8% longer than the above subject's pathway. One possible reason is that she, along with several other subjects, was sensitive to symmetrical properties of the array that eluded the elastic net. For example, several subjects reported that this array seemed to divide naturally into an upper and a lower constellation of points.

Experiment 2

MacGregor and Ormerod (1996) found virtually no correlation in subjects' performance across different arrays, and concluded that performance on visually presented TSPs appears to be determined by global perceptual properties to which the human visual system is naturally attuned. The close similarity between the performance of the two groups in the above experiment is consistent with the view that the perception of minimal structure is indeed a natural tendency of the visual system. However, the consistency in performance observed (in both groups) across different arrays calls into question the notion that performance in the TSP task is unaffected by individual differences in perceptual or cognitive abilities. It suggests instead that, provided the task is sufficiently difficult, consistent individual differences in performance will emerge. The second experiment was undertaken to examine this question more closely.

Method

Subjects. A total of 40 second-year Psychology students acted as subjects as part of a class practical exercise.

Stimulus material. A single stimulus was selected from the previous experiment. This consisted of 40 points, randomly distributed within a rectangular area, with 8 points on the convex hull. This stimulus was chosen because it showed the greatest path uncertainty, and performance on it correlated more highly and consis-

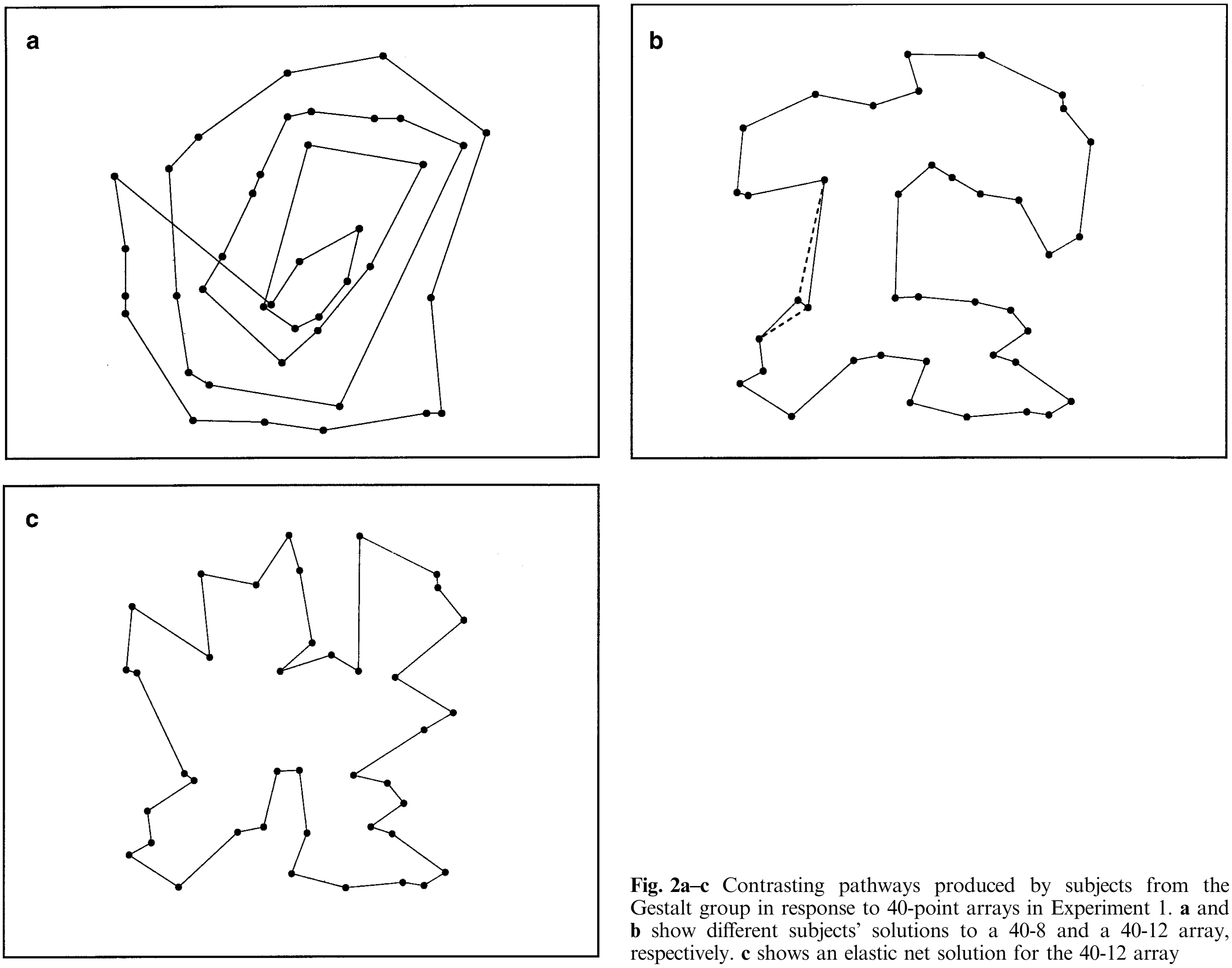

Fig. 2a–c Contrasting pathways produced by subjects from the Gestalt group in response to 40-point arrays in Experiment 1. **a** and **b** show different subjects' solutions to a 40-8 and a 40-12 array, respectively. **c** shows an elastic net solution for the 40-12 array

tently with each of the other arrays. Because path uncertainty indicates a greater variety of solution pathways, this stimulus should provide greater scope for the emergence of individual differences.

Procedure. All subjects were tested on Raven's Advanced Progressive Matrices (APM; Raven, Court, & Raven, 1988), a test in which subjects are required to complete vertical and horizontal sequences of transformations on elements of visual patterns. The APM is widely regarded as a relatively pure measure of general intelligence (Burke, 1958; Jensen, 1980), and is designed to differentiate among subjects at the upper end of the intelligence scale. Subjects were then presented with the TSP array at a separate session, and asked to connect up the points in accordance with the same instructions as were given to group O subjects in the previous experiment. Subjects were allowed to take as long as they wished, and no times were recorded.

Results

Calculation of objective measures

In addition to an estimate of pathway length, two further measures were calculated. The first was *fractal dimension* (Mandelbrot, 1983). In contrast to a straight line, with a single Euclidean dimension (length), a fractal curve tends to fill the plane to varying degrees, as measured by its fractal (or fractional) dimension. For example, by replacing a single straight line segment by a set of reduced and transformed copies of that segment, and then reiterating that procedure on each of the resultant line segments, curves can be produced that retain the same characteristics at different scales of magnification. If this procedure is continued, certain curves can be generated, which, in the limit, pass through *every* point in the plane, and hence have a fractal dimension of 2 (Peitgen, Jürgens, & Saupe, 1992). However, most curves will turn out to have a fractal dimension between 1 and 2. The perceived complexity of outline shapes has been shown by Cutting and Garvin (1987) to be closely correlated with a measure of their fractal dimension. A measure of fractal dimension might therefore provide a useful alternative to pathway length as an estimate of the efficiency of subjects' solution processes.

When the generation process is unknown, there are a number of ways of estimating the fractal dimension of an empirical curve. For example, an estimate of the box-counting dimension can be made by superimposing a grid with box size, s, on the curve, and counting the number of boxes $N(s)$ that contain any part of the curve. Successive estimates of $N(s)$ are made with progressively smaller values of s. The slope of the best fitting straight line for the plot of log $[N(s)]$ against log $(1/s)$ provides an estimate of the fractal dimension (Peitgen et al., 1992). We employed a related, but more efficient, *box-covering* algorithm (Voss, 1988), in which the number of boxes of size s required to cover the curve is calculated for a range of values of s.

The second measure calculated was that of the *Hausdorff distance* between each pathway and what appeared to be the unique optimal pathway, as calculated by the TRAVEL suite of algorithms. This provides a measure of the similarity in shape between each pathway and the optimal. The Hausdorff distance is a generalization of the notion of the distance between two points to the distance between two sets of points, A and B (Peitgen et al., 1992). It can be defined as the supremum, or least upper bound, of the distances from points in set A to points in set B and from points in set B to those in set A. To calculate it, the point in B that is furthest away from any point in A is found, and its distance, m, to the nearest point in A is measured. Similarly, the point in A that is furthest away from any point in B is found, and its distance, n, from the nearest point in B (which may be different from m) is measured. The Hausdorff distance is then the smaller of these two maximum distances, m and n. To obtain a measure of the distance between a subject's pathway and the optimum, a sufficient number of points was interpolated on straight line segments joining the nodal points, or 'cities', such that estimates of Hausdorff distance remained stable.

Individual differences in pathway length

Pathway lengths for different subjects ranged from 4.95 up to 6.41 units, with a mean of 5.46 (SD = 0.43). Scores on the APM ranged from a low score of 11, indicating poor performance, up to a high score of 35, with a mean of 25.65 (SD = 5.37).

Figure 3 shows examples of a pathway with the optimal length of 4.94 units, produced by the elastic net algorithm, together with short, intermediate, and long pathways produced by subjects. The intermediate and long pathways are characterized by pronounced 'inlets' and 'promontories', and are consequently much less convex than the optimal or near-optimal solutions.

Correlations among dependent measures

Table 3 shows the Spearman rank order correlations between the various pairs of measures. The length, fractal dimension, and Hausdorff distance from the optimal pathway all show significant correlations with performance on Raven's APM, as well as with each other.

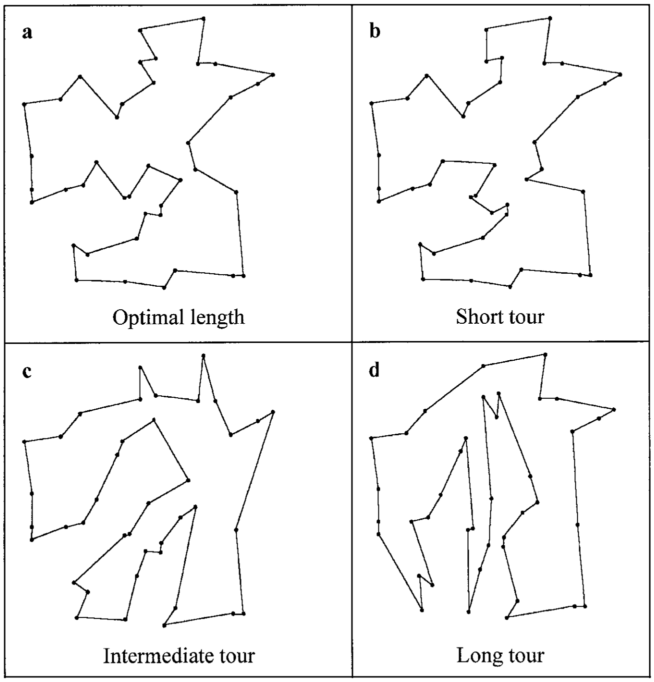

Fig. 3 Examples of **a** the optimal pathway, and **b** a short, **c** an intermediate, and **d** a long pathway produced in Experiment 2

Table 3 Spearman correlation coefficients for TSP path length (TSP), fractal dimension (Fractal D), Hausdorff distance (Hausdorff) and APM scores (APM) (*TSP* Traveling Salesman problem, *APM* advanced progressive matrices)

	TSP	Fractal D	Hausdorff	APM
TSP	—	0.89**	0.69**	−0.36
Fractal D		—	0.60**	−0.34
Hausdorff			—	−0.50*
APM				—

Note: all correlations are one tailed with $N = 40$
$^{*}p < 0.01$; $^{**}p < 0.001$

Discussion

The above results can be interpreted at both a specific and a general level. At the specific level, Experiment 1 showed that, in the case of each group, both the total number of points and the number of points on the convex hull had significant effects on the speed of pathway completion. Irrespective of the task, subjects produced complete pathways more quickly when there were more points on the convex hull, and this advantage was (generally) more pronounced when the arrays contained a higher total number of points. The effects of both total number of points and of convexity are echoed in the lower path uncertainty scores, observed in group O, for those arrays with a smaller total number of points and with more points on the convex hull. At the same time, there was no evidence, in either group, that the standardized deviations of subjects' pathways from the optimum were influenced by either the total number of points in an array or by its degree of convexity.

These findings mean that the primary conclusion of MacGregor and Ormerod (1996) is subject to considerable qualification. Within the parameter ranges investigated here, neither total number of points nor convexity affect the optimality of subjects' pathways, whatever the instructions. However, despite the strictures of Lee and Vickers (2000), convexity does appear to facilitate the perception of Gestalt structures, which usually happen to be both 'good' and near-optimal, and this facilitation appears to be greater as the number of points on the convex hull is increased.

Experiment 1 also showed that subjects under Gestalt instructions produced completed pathways more quickly than those under instructions to find a minimal pathway, and that their pathways were somewhat longer on average. At this stage, our best guess is that these differences may reflect the results of more local, cognitively guided processes in the O group, as well as the increased time required by such processes. An additional factor, which may also have contributed to the greater variance in path length shown by the G group, is that some subjects in this group may have looked for pathways that maximized some form of symmetry (such as spiral), even though this did not minimize path length. At the same time, the extent to which the G group focussed on producing minimal solutions is remarkable. As shown in Table 1, the three shortest solutions frequently reoccurred in the G group, and did so almost as often as they reoccurred in the O group. These findings are certainly consistent with MacGregor and Ormerod's conclusion that the perception of minimal structure may be a natural, automatic tendency of the human visual system.

Also consistent with this conclusion are reports by Polivanova's subjects that they looked for 'natural' forms. This description would apply, for example, to elastic net solutions to problems with several hundred nodes, which have an attractive, fractal appearance, reminiscent of branching coral (e.g., Fritz & Wilke, 1991). Such structures are frequently the signature of a naturally occurring recursive process that is itself attempting to minimize (or maximize) the value of some function under a multiplicity of constraints (Mandelbrot, 1983; Barnsley, 1988). It may be that subjects are sensitive to the underlying order that results from such optimizing, and this is why they find the structures attractive (Pickover, 1990; Schmidhuber, 1997). It is also possible that subjects' attempts to find optimal pathways, by constructing natural, aesthetically pleasing structures, reflect a preference for capturing maximally disordered data by means of an algorithm of minimum complexity (Chaitin, 1987). These two alternatives are closely related, and it may be possible to find some more general formulation that would encompass them both. In any case, it seems likely that the detection of approximate, partial symmetries is an important ingredient in the process. This would go some way towards explaining the marked differences in length found between some pathways, such as those shown in Fig. 4a and b, which are nevertheless both characterized by a high degree of apparent symmetry. However, the possible relationship between symmetry maximization and the minimization of length (or, possibly, perimeter) in the spontaneous perception of organization in such random arrays is still far from clear.

Meanwhile, and contrary to the previous data, human performance on visually presented TSPs appears to show consistent individual differences that are reliable across different problems, and are related to performance in a psychometric test of general intelligence. Although the measures of pathway length, fractal dimension and Hausdorff distance are intercorrelated, it is arguable that they each capture a different aspect of performance: namely, minimality, complexity, and similarity in shape to the optimal pathway. Only further research will tell which measure(s) will turn out to be the most informative.

The fact that a number of subjects in both conditions outperformed the elastic net rules out this algorithm as an account of the perceptual optimization process. Since the performance of the elastic net declined at a faster rate than that of the subjects, as the size of the arrays increased, this algorithm proved to be an insufficient model of the subject's approach to the TSP. Instead, one possibility is that, under these conditions, subjects become increasingly sensitized to symmetries in the arrays. These trigger Gestalt organizations which, in the case of group O, may be modified by attention to more local features. There is considerable evidence that subjects are sensitive to a wide variety of approximate and incomplete symmetries (e.g., Tyler, 1996), and one of our avenues for future research is directed towards the development of a general algorithm for the detection of such symmetries.

On a more positive note, the elastic net did demonstrate the relative success of the human TSP solving style (whatever that may be). In addition, the individual differences in performance suggest that the particular problem solving approach involved may be more complex than the process encapsulated by this model. Although many subjects confirmed the results of previous research by claiming that they utilized global, geometrical properties of the arrays, it appeared that the problem was subsequently broken down into a number of smaller tasks involving sub-groups of points. We are currently developing a modified elastic net algorithm, in which the initial ring of nodes takes the form of the convex hull of the array, and is subsequently deformed until it encompasses the remaining points in the problem. Such an algorithm would take into account the 'two-stage' model alluded to above, as well as being sensitive to the composition of the convex hull.

Finally, because there is no difficulty in generating multiple, stochastically equivalent arrays, the consistent individual differences, found in both experiments, suggest that the visual TSP task may have some potential as a neuropsychological test (Geffen, 1995), and that optimization might provide a useful perspective in which to investigate path-following performance more generally. Meanwhile, the link with intelligence, and the occurrence of optimal structures in the natural world, suggest that the perception of optimal structure may have some adaptive utility. This suggests that the type of task studied here may not only be interesting from the perspective of perception, and cognition, but may help to provide a conceptual framework of optimization, within which to study intelligence in general.

Such a perspective seems to have much in common with Anderson's (1990; 1991) adaptive theory of cognition. Anderson (1991) argues that there is a serious and perhaps intractable induction problem in inferring the structure of mental processes from that of the behavior they produce. There is also a serious and possibly intractable identifiability problem, in that many different mechanisms may give rise to similar patterns of behavior. Anderson goes on to argue that the hypothesis that human cognition constitutes an optimal response to the information-processing demands of the environment provides a useful constraint on the kinds of mental mechanism we might propose. This argument is similar to Marr's (1982) view that an understanding of algorithmic procedures is incomplete without some insight into what the algorithm in question has been designed (or has evolved) to achieve. A similar argument may apply to the study of human intelligence. At the very least, an optimization perspective suggests what intelligence might be *for*.

Acknowledgements A substantial part of the material in this paper is based on an unpublished thesis, submitted by M. A. Butavicius for the degree of B.A. (Honors Psychology) in the University of Adelaide. The research in this paper benefited from a University Research Grant to D. Vickers.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences*, *14*, 471–517.

Anderson, M. (1992). *Intelligence and development*: *a cognitive theory*. Oxford: Blackwell.

Barnsley, M. (1988). *Fractals everywhere*. New York: Academic Press.

Boyd, S. C., Pulleyblank, W. R., & Cornuéjols, G. (1987). TRAVEL – An interactive traveling salesman problem package for the IBM-personal computer. *Operations Research Letters*, *6*, 141–143.

Burke, H. R. (1958). Raven's progressive matrices: a review and critical evaluation. *Journal of Genetic Psychology*, *93*, 199–228.

Butler, M., Retzlaff, P. D., & Vanderploeg, R. (1991). Neuropsychological test usage. *Professional Psychology*: *Theory, Research, and Practice, 22*, 510–512.

Chaitin, G. J. (1987). *Algorithmic information theory*. Cambridge, UK: Cambridge University Press.

Cutting, J. E., & Garvin, J. J. (1987). Fractal curves and complexity. *Perception and Psychophysics*, *42*, 365–370.

Durbin, R., & Willshaw, D. (1987). An analogue approach to the Traveling Salesman problem using an elastic net method. *Nature, 326*, 689–691.

Durbin, R., Szelski, R., & Yuille, A. (1989). An analysis of the elastic net approach to the Traveling Salesman problem. *Neural Computation*, *1*, 348–358.

Flood, M. M. (1956). The traveling salesman problem. *Operations Research*, *4*, 61–75.

Fritzke, B., & Wilke, P. (1991). FLEXMAP: A neural network for the traveling salesman problem with linear time and space complexity. *International Joint Conference on Neural Networks (Singapore)*, 929–934.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist*, *58*, 34–42.

Geffen, G. (1995). Approaches to neuropsychological assessment: A comment. *Australian Psychologist*, *30*, 45–46.

Glass, L. (1969). Moiré effects from random dots. *Nature*, *223*, 578–580.

Golden, B., Bodin, L., Doyle, T., & Stewart, W. (1980). Approximate traveling salesman algorithms. *Operations Research*, *28*, 694–711.

Goldschlager, L., & Lister, A. (1988). *Computer science*: *a modern introduction*. New York: Prentice Hall.

Hildebrandt, S., & Tromba, A. (1996). *The parsimonious universe*: *shape and form in the natural world*. New York: Copernicus.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Lawler, E. L., Lenstra, J. K., Rinooy Kan, A. H. G., & Shmoys, D. B. (1985). *The traveling salesman problem*: *a guided tour of combinatorial optimization*. Chichester, UK: Wiley.

Lee, M. D., Brown, M., & Vickers. D. (1997). Neural network and tree search algorithms for the generation of path-following (trail-making) tests. *Journal of Intelligent Systems*, *7*, 117–143.

Lee, M. D., & Vickers, D. (2000). The importance of the convex hull for human performance on the traveling salesman problem: a comment on MacGregor and Ormerod (1996). *Perception and Psychophysics*, *62*, 226–228.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lin, S., & Kernighan, B. W. (1973). An effective heuristic for the traveling salesman problem. *Operations Research*, *21*, 498–516.

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

Malsburg, C. von der, & Willshaw, D. J. (1977). How to label nerve cells so that they can interconnect in an ordered fashion. *Proceedings of the National Academy of Science*, *USA*, *74*, 5176–5178.

Mandelbrot, B. B. (1983). *The fractal geometry of nature*. New York: Freeman.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Norback, J. P., & Love, R. F. (1977). Geometric approaches to solving the traveling salesman problem. *Management Science*, *23*, 1208–1223.

Oswald, W. D., & Roth, E. (1987). *Der Zahlen-Verbindungs-Test (ZVT) Handanweisung (Manual)*. Göttingen, Germany: Hofgrefe.

Peitgen, H.-O., Jürgens, H., & Saupe, D. (1992). *Chaos and fractals: new frontiers of science*. New York: Springer.

Peterson, C. (1990). Parallel distributed approaches to combinatorial optimization: Benchmark studies on the traveling salesman problem. *Neural Computation*, *2*, 261–269.

Peterson, C., & Söderberg, B. (1995). Neural optimization. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 617–621). Cambridge, MA: MIT Press.

Pickover, C. (1984). The use of random dot displays in the study of biomolecular conformation. *Journal of Molecular Graphics*, *2*, 34.

Pickover, C. (1990). *Computers, pattern, chaos and beauty: graphics from an unseen world*. New York: St. Martin's Press.

Polivanova, N. I. (1974). [Functional and structural aspects of the visual components of intuition in problem solving]. *Voprosy Psikhologii, 4,* 41–51.

Raven, J. C., Court, J. H., & Raven, J. (1988). *Manual for Raven's progressive matrices and vocabulary scales. Section 4: Advanced progressive matrices*. London: Lewis.

Reitan, R. M. (1992). *Trail Making test: manual for administration and scoring*. Tucson, AZ: Reitan Neuropsychology Laboratory.

Schmidhuber, J. (1997). Low-complexity art. *Leonardo*, *30*, 97–103.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Shepard, R. N. (1984). Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, *91*, 417–447.

Simmen, M. W. (1991). Parameter sensitivity of the elastic net approach to the traveling salesman problem. *Neural Computation*, *3*, 363–374.

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

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.

Tyler, C. W. (1996). *Human symmetry perception and its computational analysis*. Utrecht, The Netherlands: VSP.

Vickers, D., & Lee, M. (1998). Never cross the path of a traveling salesman: the neural network generation of Halstead-Reitan Trail Making tests. *Behavior Research Methods*, *Instruments*, *& Computers*, *30*, 423–431.

Voss, R. F. (1988). Fractals in nature: from characterization to simulation. In: H.-O. Peitgen, & D. Saupe (Eds.). *The science of fractal images* (pp. 21–70). New York: Springer.

Willshaw, D. J., & Malsburg, C. von der (1979) A marker induction mechanism for the establishment of ordered neural mappings; its application to the retinotectal problem. *Philosophical Transactions of the Royal Society*, *B287*, 203–243.

Yuille, A. L. (1995). Constrained optimization and the elastic net. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 250–255). Cambridge, MA: MIT Press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison Wesley.