

# Human performance on the traveling salesman problem

J. N. MACGREGOR and T. ORMEROD

Loughborough University of Technology, Loughborough, England

Two experiments on performance on the traveling salesman problem (TSP) are reported. The TSP consists of finding the shortest path through a set of points, returning to the origin. It appears to be an intransigent mathematical problem, and heuristics have been developed to find approximate solutions. The first experiment used 10-point, the second, 20-point problems. The experiments tested the hypothesis that complexity of TSPs is a function of number of nonboundary points, not total number of points. Both experiments supported the hypothesis. The experiments provided information on the quality of subjects' solutions. Their solutions clustered close to the best known solutions, were an order of magnitude better than solutions produced by three well-known heuristics, and on average fell beyond the 99.9th percentile in the distribution of random solutions. The solution process appeared to be perceptually based.

The Euclidean version of the traveling salesman problem (TSP) consists of finding the shortest closed path (tour) through a set of points in the plane. That is, the shortest path must be found which passes through every point and returns to the origin. The problem has become a classic in the field of combinatoric optimization and has accumulated a vast body of research in the fields of management science and operations research. Complexity theory classifies the problem as "hard" and as "NP-complete." The distinction between *easy* and *hard* problems is based on the time that it would take an algorithm to solve a problem with  $n$  objects. A problem is "easy" if there exists an algorithm that can solve any instance of it in polynomial time (i.e., in a time proportional to  $n^k$ , where  $k$  is some integer). The TSP is "hard" because the time that even the most efficient known algorithms require in order to solve it grows exponentially as a power of  $n$ . By way of comparison, an algorithm solving in time  $n^3$  would take about .001 sec for a problem with 10 objects, operating at a rate of one instruction per microsecond. An algorithm operating in time  $2^n$  would take roughly the same time. But for a problem of 100 objects, the former would take about 1.0 sec, whereas the latter would need  $10^{15}$  centuries (Barton, Berwick, & Ristad, 1987).

A problem is classed as NP, or nondeterministic polynomial, if there are nondeterministic algorithms which may solve it in polynomial time. Such algorithms may be thought of as operating in two stages—in the first, guessing at a solution, and in the second, checking whether or not the solution is correct. Roughly speaking, NP is the class of problems in which correct or optimal solutions

are hard to find but easy to check. NP problems classified as "complete" have the additional important property that each NP-complete problem is quickly reducible to every other NP-complete problem. This means that if an algorithm can be found that solves any one NP-complete problem in polynomial time, then there exists a polynomial time algorithm for every NP-complete problem, and there are no "hard" problems. Alternatively, if it can be proved that there is no polynomial time algorithm for one NP-complete problem, then there is none for any of them, and all will have been proved to be intransigent. Hundreds of problems have now been proved to be NP-complete, including the TSP. It has been suggested that the question of whether or not there exist polynomial time algorithms for solving NP-complete problems is one of the most important unsolved problems in theoretical computer science (Wilf, 1986).

In addition to its theoretical interest, the TSP has practical implications for problems as apparently diverse as circuit-board drilling, X-ray crystallography, and the laying of ducting (Krolak, Felts, & Marble, 1971; Sangalli, 1992). Some practical problems that involve circuit-board drilling, or laser movements in chip manufacture, translate into TSPs of up to a million nodes (Sangalli, 1992). However, as noted above, no practical algorithm has been discovered that is certain to find the optimal solution to TSPs, and a guaranteed method—exhaustive search—becomes increasingly impractical as the number of points ( $n$ ) increases. The number of possible solutions to a TSP increases by  $(n - 1)!/2$  (if the direction of the tour is ignored), so that in order to find all possible solutions to a 20-node problem, a computer generating, say, 100 solutions per second would require in excess of 19 million years. The TSP therefore has a number of intriguing characteristics: it is easy to state, it is easy to represent, but it defies an economical general method of solution. This last factor has stimulated interest in developing

T.O. is now in the Department of Psychology at the University of Lancaster, Lancaster, England. Correspondence concerning this article should be addressed to J. N. MacGregor, School of Public Administration, University of Victoria, P.O. Box 1700, Victoria, BC, Canada V8W 2Y2 (e-mail: jmacgreg@hds.uvic.ca or t.ormerod@lancaster.ac.uk).

approximate heuristic procedures, and many ingenious procedures have been designed and compared. In order to find "reasonable" solutions, such heuristic procedures generally need to perform in the order of  $n^3$  calculations. An approximate solution of a 20-node problem would therefore require some 8,000 calculations (Golden, Bodin, Doyle, & Stewart, 1980).

Given the characteristics of the problem, it would be of interest to know how well human subjects perform on TSPs. However, although the TSP can be presented easily to human subjects, it appears to have received little attention as a topic of psychological research. Our literature search failed to identify any published articles that directly addressed human performance on the problem. The only published studies to our knowledge that used human participants did so in an effort to improve machine solutions, not to investigate human performance per se (Krolak et al., 1971; Michie, Fleming, & Oldfield, 1968). The former study, for example, used computer procedures to initially organize the data into clusters. The clusters were then presented to human operators for connection into a completed tour. The results suggested that humans may have been quite skilled in their role, in that the human-computer procedure produced high-quality solutions in relatively short times (Krolak et al., 1971).

It appeared to us that the TSP might provide an interesting psychological task to explore. It seems that it may be fundamentally a perceptual problem-solving task, involving as it does the construction of a visually represented solution from a visually represented problem space. There is also the suggestion from the Krolak et al. (1971) results that humans may be capable of achieving reasonable solutions with minimal cognitive load. The purpose of the present research was, therefore, to collect some preliminary information on human performance, and to investigate some factors that might affect the complexity of TSPs for human subjects.

It might initially appear that the complexity of TSPs would simply be a function of the number of points to be connected. However, an early proof about the optimal solution suggests that this might not be the case. The points making up a TSP can be considered as a convex set (the convex hull), and it has been shown that the optimal path for every Euclidean TSP connects adjacent points on the boundary of the convex hull in sequence, though the path may pass through interior points between adjacent boundary points (Flood, 1956). Figure 1 illustrates a solution which violates this principle (bottom) and one which conforms to it (top). Clearly a mechanism that finds optimal paths must adhere to this principle, and for such a mechanism there would be no uncertainty about the order in which to connect boundary points, regardless of how many there were. If human performance followed the same principle, at least approximately, then the complexity of a TSP should depend not so much on the total number of points, but on the number of nonboundary, or interior, points. The first experiment was designed to test this hypothesis.

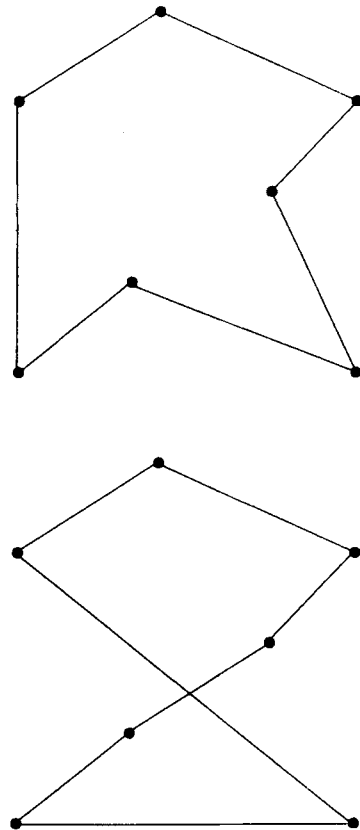


Figure 1. Two possible solutions to a Traveling Salesman Problem. The bottom solution violates a principle which has to be met in finding the optimal solution.

### Measurement Issues

In addition to testing the preceding hypothesis, the experiment was designed to collect information on the quality of human solutions. This raised certain measurement issues. One apparent measure of solution quality is the length of the path, but this is unsuitable for comparing performances across problems, since path lengths may vary depending on the geometry of specific problems. One of the basic measures that appears in the operations research literature for comparing the performances of algorithms takes the difference between a given solution path length and the optimal path length, expressed as a percentage of the optimal path length. If the optimal is unknown, the best known solution is used as a surrogate (Golden et al., 1980). This measure is suitable for comparing different algorithms on the same benchmark problems, but it may be less suitable for comparing performances across problems, because it makes no allowance for variation in the density or sparsity of solutions close to the optimal for different problems.

A second approach, more familiar to psychology, is to express subjects' path lengths as standard scores based on the mean and standard deviation of the population of all possible path lengths, estimated from a random sample. This approach also has difficulties. It assumes that

the parent population approximates a normal distribution, whereas we know that it will differ from normality in some specific respects. It will not be continuous, for example, though the very large number of solutions may ameliorate this concern. It will also have a strict upper and lower bound. Generally, the shape of the population distribution will be unknown.

To get some idea of the quality of human solutions, both of these imperfect approaches were adopted here. In addition, the results of the experimental subjects were compared with solutions generated by three heuristic procedures. The heuristics used were the Nearest Neighbor algorithm, the Largest Interior Angle algorithm (Norback & Love, 1977), and a variation of a Convex Hull heuristic described by Golden et al. (1980), using a cheapest insertion criterion (CHCI). These three were used because their rules might be employed by human subjects and might conceivably form the basis of a model of human performance. In addition, some Convex Hull-based approaches have been found to produce optimal or near-optimal solutions with a high degree of reliability.

The algorithms used the following rules: For the Nearest Neighbor method, an initial point is selected as the start of the path. Define this as the leading point. The unconnected point closest to the leading point is then added to the path, and it becomes the leading point. This is repeated until all points have been added, at which time the leading point is connected to the starting point to complete the tour. Both the Largest Angle and the Convex Hull methods begin by connecting the points on the boundary of the hull into an initial subtour. An informal analogue of these procedures is to think of the points as pegs in a board and the path as a rubber band stretched around the pegs. Initially, then, the band will touch all pegs on the boundary of the convex hull. Next, the segment of the band which is "closest" to an unconnected peg is then stretched in to incorporate that peg. This is repeated until the band winds around all of the pegs. The two heuristics differ in how they identify the "closest" unconnected point. With the Largest Angle method, after the boundary points have been connected into an initial subtour, the points  $i, j$  in the subtour and  $k$  not in the tour are found for which the angle  $i, k, j$  is maximal, and  $k$  is inserted between  $i$  and  $j$ . This step is repeated until all points are in the tour. With the version of the Convex Hull heuristic used here, the insertion rule was to find the  $i, j, k$ , for which the distance  $(i, k) + (k, j) - (i, j)$  is minimal, and insert  $k$  between  $i$  and  $j$  (cheapest insertion). This is repeated until all points are in the tour.

It will be noted that the Nearest Neighbor algorithm can produce different tours, depending on which point is selected as the starting point. Both the Largest Angle and the Convex Hull approaches, on the other hand, typically yield a single tour. (There is a minor exception to this: If 2 or more points are tied for insertion at the same stage in a solution, then pursuing all possible solution options may produce several different paths.)

## EXPERIMENT 1

### Method

#### Subjects

The 58 subjects were students at Loughborough University of Technology.

#### Materials

Six TSP problems with 10 points were generated, with the number of interior points varying from 1 to 6. The problems were created in the following way: The coordinates of the points on the boundary of the hull were initially determined from the vertices of the regular polygon with that number of sides, using polar coordinates with a radius of 80 mm. For example, for 4 points on the boundary of the hull, the coordinates were determined from the interior angles  $45^\circ$ ,  $135^\circ$ ,  $225^\circ$ , and  $315^\circ$ . Some irregularity was then introduced by adding to each angle a value selected randomly from the range of  $+5^\circ$  to  $-5^\circ$ . Once the points on the boundary of the hull had been determined, polar coordinates for the interior points were generated randomly with the constraint that they fell within 40 mm of the center of the original regular polygon. Once 10 points had been generated for each problem, the polar coordinates were converted to rectangular coordinates centered on the coordinates (140, 90). These will be referred to as the *basic problems*. A seventh 10-point problem, taken from the operations research literature, was also included (Dantzig, Fulkerson, & Johnson, 1959). This *Dantzig* problem was used because it represents a benchmark test, one that many relatively effective heuristics fail to solve. The problem had 5 interior points.

The problems were drawn on separate  $11 \times 8.5$  in. sheets. Figure 2 shows the problem with 6 interior points on the bottom and the Dantzig problem on the top, with their optimal paths. The coordinates for all of the problems are given in the Appendix.

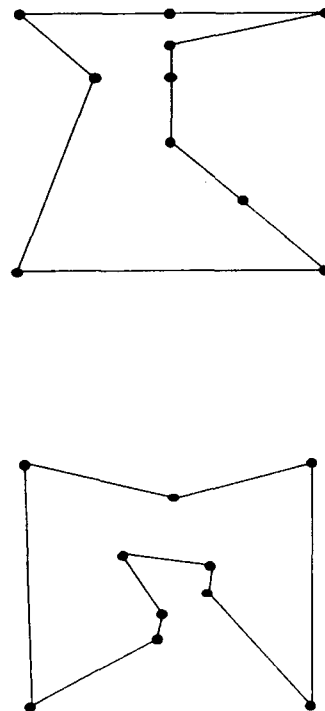


Figure 2. Two of the problems used in Experiment 1 with the points connected by the shortest paths.

### Design

Each subject was tested on each problem. Three different orders of presentation were used. In one, the problems were presented in the order 6, 5, 4, . . . , 1 interior points. In a second, this order was reversed. In the third, problems were presented in the random order 3, 5, 2, 6, 1, and 4 interior points. In each case the Dantzig problem was presented last.

### Procedure

Data was collected in a classroom setting. Each subject was given a booklet of problems with a single page of instructions. The instructions indicated that for each problem the subjects should select a starting point and then draw what they thought to be the shortest path from the starting point, passing through each point and returning to the starting point. Subjects were asked to indicate the starting point and direction of travel for each problem. No time limit was imposed, but the instructions indicated that a maximum of 5 min should be taken per problem. Subjects handed in their booklets when finished, and times were recorded to the nearest minute.

### Results

Ten subjects left one or more of the problems incomplete by failing to incorporate 1 or more points, and these cases were excluded from analysis. A further 3 completed protocols were randomly eliminated to leave a total of 45 cases, 15 from each order. Path lengths for subjects' solutions were obtained and provided the primary data for analysis. Times to complete the seven problems varied from 6 to 25 min, with an average of 13 min. There were no effects of order on either path length or solution times, and the results were collapsed across the three orders.

### Quality of Solutions

**Standard score approach.** For each of the basic problems a random sample of 100 solutions was generated, and the sample means and standard deviations were used to convert subjects' path lengths to standard scores. The main results appear in Table 1, which shows, for each problem, the minimum, mean, and maximum path lengths observed over the 45 subjects. Conversion of the standard scores to percentiles shows that minimum path lengths were on the average beyond the 99.99th percentile; mean path lengths, beyond the 99.97th percentile; and maximum path lengths, beyond the 99.45th percentile. By this reckoning, average best and worst performances were within roughly half a percentile point of each other. Of the 270 solutions to the basic problems, the very worst was

at the 97th percentile. The results for the Dantzig problem are shown at the foot of Table 1. The problem had 5 interior points and was drawn to a larger scale than the others, and consequently produced much longer path lengths. The results are by and large consistent with those of the basic problems, with percentiles of 99.94, 99.84 and 97.86, for minimum, mean, and maximum path lengths, respectively.

**Percentage over optimal approach.** The three heuristics described previously were used with each of the problems. For the Nearest Neighbor heuristic, all possible starting points were used. The optimal solutions for the six basic problems were found by using TRAVEL, a system of programs designed to produce provably good solutions to TSPs (Boyd, Pulleyblank, & Cornuéjols, 1987). TRAVEL can identify the lower bound of an optimal solution, so that if a solution is found that equals the lower bound, it must be the optimal solution. For the Dantzig problem, the optimal solution was already known.

The path lengths produced by the subjects and the heuristics for each problem were converted to a percentage over the optimum. The main results are shown in Table 2. The table shows the percentage over the optimum of the best solution produced by the subjects and each of the three heuristics. It will be recalled that the CHCI and Largest Angle procedures produce only one solution per problem, while the human subjects and Nearest Neighbor procedure can produce a variety of solutions. To reflect this, the table also shows the percentage over the optimum for the subjects' mean and the Nearest Neighbor mean (the latter was the average of the path lengths produced by using every possible starting point).

### Problem Complexity

Following the principle that good patterns have few alternatives (Garner, 1970) a first approximation to a measure of the difficulty of problems is the variety of different solutions generated. The numbers of different subject solutions for the problems were 3, 10, 12, 18, 15, and 25 for the problems with 1, 2, 3, 4, 5, and 6 interior points, respectively. The Dantzig problem, with 5 interior points, generated 20 different solutions. The correlation between number of interior points and number of different solutions was  $r = .94$  ( $p < .01$ ). The number of different solutions may be a somewhat crude measure, in that it does not distinguish between solutions that differ slightly and those that differ radically. For this reason, a second measure was used, based on the response uncertainty associated with all possible pairwise connections between the 10 points of a problem. For each problem, the number of subjects connecting each pair of points was tabulated, and the resulting frequencies were converted to probabilities. These were used to calculate the uncertainty in bits associated with each connection, which were then summed to provide total response uncertainty for each problem. To illustrate the calculation, imagine that 3 subjects produced paths for a 5-node problem, with 2 subjects connecting the nodes in the sequence 1, 2, 3, 4, 5, 1, the third in the sequence 1, 2, 3, 5, 4, 1. The re-

**Table 1**  
Minimum, Mean, and Maximum Path Lengths Produced by Subjects in Experiment 1, and the Corresponding *z* Values

Number of Interior Points	Subjects' Path Lengths					
	Minimum		Mean		Maximum	
	Score	<i>z</i>	Score	<i>z</i>	Score	<i>z</i>
1	530.89	4.75	535.14	4.71	582.32	4.31
2	566.22	4.51	586.11	4.32	697.14	3.27
3	559.55	4.12	574.99	3.95	647.73	3.15
4	595.30	4.32	617.94	3.99	665.00	3.31
5	558.15	4.08	575.15	3.80	638.52	2.77
6	528.51	3.49	548.83	3.11	613.33	1.93
Dantzig	758.66	3.45	787.21	3.15	870.24	2.29

**Table 2**  
**Optimal Path Length and Percentage Above the Optimal for Human and Heuristic Solutions, Experiment 1**

Number of Interior Points	Optimal	Percentage Above Optimal					
		Subject Minimum	NN Minimum	LA	CHCI	Subject Mean	NN Mean
1	530.89	0	0.0	0.0	0.0	0.8	0.0
2	566.22	0	0.0	0.0	0.0	3.5	0.0
3	559.55	0	0.0	5.3	2.1	2.8	6.6
4	595.30	0	0.0	11.5	2.4	3.8	10.0
5	558.15	0	4.0	6.1	1.0	3.0	9.1
6	528.51	0	5.9	2.9	0.0	3.8	10.0
Dantzig	758.66	0	0.6	3.0	2.7	3.8	10.4

Note—NN, Nearest Neighbor; LA, Largest Internal Angle; CHCI, Convex Hull, with cheapest insertion criterion.

sulting frequency matrix would indicate that the points 1 and 2 were connected by all 3 subjects, as were the points 2 and 3, and 4 and 5, resulting in probabilities of 1.0 for these connections. The points 1 and 5 were connected by 2 subjects, as were the points 3 and 4, yielding probability values of .67 for these connections. Finally, the connections 1 and 4, and 3 and 5, were made only once, giving probabilities of .33. The total uncertainty, *H*, associated with the connections would then be found by entering the observed probabilities into the Shannon formula,

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

where *k* is the total number of connections made by subjects.

Figure 3 (solid squares) shows the response uncertainty for each problem plotted as a function of the number of interior points. The product-moment correlation between uncertainty and the number of interior points for the six basic problems was *r* = .93 (*df* = 4). The relationship was significant beyond the .01 level (two-tail test).

(Adding the Dantzig problem, which had a response uncertainty of 9.89 bits, did not substantially alter the results, and the correlation remained at *r* = .93, *df* = 5).

**Other Findings**

**Individual differences.** To test whether or not individuals systematically differed in their ability to solve the problems, the path lengths for the subjects' solutions to each of the six basic problems were ranked and the average Spearman rank-order correlation between each pair of rankings was obtained. The resulting average was .06 (*n* = 45), indicating virtually no correlation in subjects' performances across different problems.

**Order of boundary points.** The optimal path to a TSP connects adjacent boundary points in order. The large majority of subjects' solutions displayed the same characteristic. Three hundred and ten of the 315 solutions (98%) adhered to this principle. The 5 exceptions were produced by 5 different subjects, 3 in the case of the Dantzig problem and 1 each in the problems with 2 and 6 interior points.

**Crossed arcs.** Failure to connect boundary points in order of adjacency automatically creates a solution with

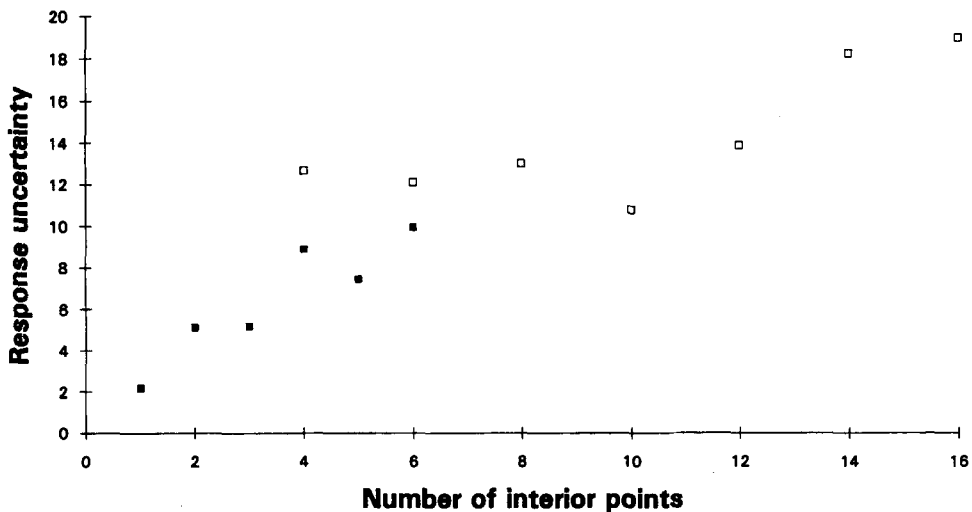


Figure 3. The response uncertainty to line connections as a function of the number of interior points. (Closed squares show the results for the 10-point problems; open squares, for the 20-point.)

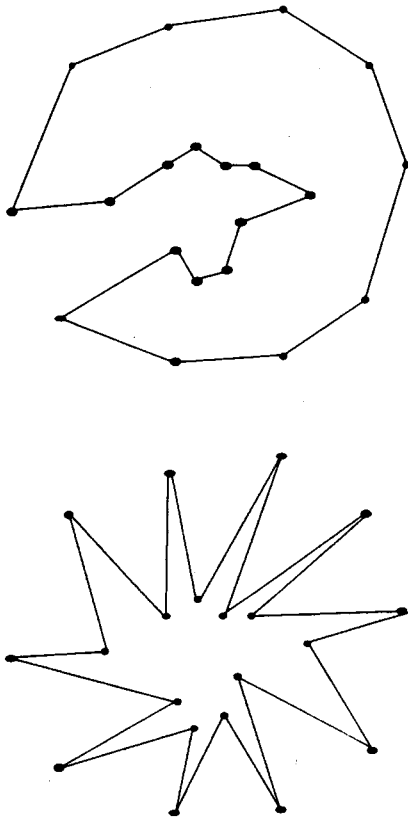


Figure 4. Two different solutions to the same 20-point problem, the one on the top with the minimum number of indents, the one on the bottom with the maximum.

crossed arcs. However, crossed arcs can also occur when boundary points are connected in order. There were no such additional cases of crossed lines.

**Number of indentations.** If subjects are guided by the global properties of the convex hull, they might exhibit a preference not only for connecting points on the boundary in sequence, but also for connecting adjacent points on the boundary to *each other*. Any such preference would tend to result in solutions with fewer rather than more *indentations*. An indentation occurs whenever an internal point is connected to a boundary point, and solutions may vary in the number of indentations they exhibit. It is possible to complete any TSP with only one indentation, provided there are at least 3 boundary points and 1 interior point. On the other hand, it is possible to

have as many indentations as there are boundary points, provided there is a sufficient number of interior points. Figure 4 illustrates two different solution paths to the same problem, one with the maximum number of indentations, the other with the minimum. Are subjects' solutions more like the figure on the top than would be expected on the basis of the availability of indents in a problem? The maximum number of indentations that a problem can have is determined by the number of boundary points or the number of interior points, whichever is less. In other words, if  $n$  is the number of points, and  $m$  is the number of interior points, then the number of indentations can vary from 1 to either  $m$  or  $n - m$ , whichever is less. The maximum indentations for the problems with 1–6 interior points were therefore 1, 2, 3, 4, 5, and 4, respectively. If subjects have no particular preference for connecting boundary points to each other, we would expect the number of indentations produced to be proportional to their availability within problems. If there were such a preference, we would expect subjects' solutions to have fewer indentations than this. The relevant results are shown in Table 3. The first column gives the number of possible indentations for five of the six basic problems plus the Dantzig problem. (Since there are no degrees of freedom for the problem with 1 interior point, it has been omitted.) The second column shows the number of problems for which that number of indentations is possible. Thus all six allowed 1 and 2 indentations, five allowed 3, and so on. Only two problems allowed as many as 5 indentations. Each of these frequencies was then divided by the column total, to give the relative frequencies shown in the third column, which are then expressed as a cumulative frequency distribution in column 4. Columns 5, 6 and 7 show the corresponding empirical results. Column 5 shows the raw frequencies of each solution type; column 6, the relative frequencies; and column 7, the cumulative frequency distribution. Comparison of entries from column 7 with those from column 4 shows that 84% of subjects' solutions had two or fewer indentations, compared with an expected value of 52%, based on the availability of this type of solution. The largest absolute difference between corresponding values from columns 4 and 7,  $D_{MAX}$ , provides the statistic for the Kolmogorov–Smirnov test to test the significance of the differences between the observed and expected distributions. The results indicated that subjects' solutions had significantly fewer indentations than was

Table 3  
Expected and Observed Frequency, Relative Frequency, and Cumulative Frequency of Indentations in Problem Solutions, Experiment 1

Number of Indents	Frequencies					
	Expected			Observed		
	Raw	Relative	Cumulative	Raw	Relative	Cumulative
1	6	.26	.26	91	.34	.34
2	6	.26	.52	134	.50	.84
3	5	.22	.74	39	.14	.98
4	4	.17	.91	6	.02	1.00
5	2	.09	1.00	6	.00	1.00

to be expected on the basis of availability ( $D_{MAX} = 0.32$ ,  $p < .01$ ).

### Discussion

The experiment was designed to provide initial information about the quality of human solutions to the TSP and to test the hypothesis that problem complexity is influenced by the number of internal points rather than by the total number of points making up a problem.

With respect to the quality of human solutions, the present results suggest that they are of a high quality. Comparison with a random sample indicated that the present solutions formed a tight cluster in the upper tail of the distribution. Mean performance fell beyond the 99.9th percentile, and on average, only half a percentile separated the best and worst performances. These results could be misleading, because we cannot be certain about the shapes of the distributions of the solution populations. For this reason the results were examined in other ways. Three heuristic procedures were used to generate solutions, and the best human solutions were found to compare very favorably with the best solutions found by a heuristic. For each of the six basic problems, at least 1 subject found the optimal solution, whereas no heuristic found the optimal for every case. The Nearest Neighbor procedure found the best solution for four of the six problems; the CHCI procedure, for three; and the Largest Angle procedure, for two. Averaged across the six problems, the three heuristics produced solutions which were 1.65%, 1.10%, and 4.30% above the optimal, for the Nearest Neighbor, CHCI and Largest Angle procedures, respectively. The comparable human figure was 0.00%. Similar results were obtained for the benchmark Dantzig problem. None of the heuristic procedures found the optimal solution, while 3 of the human subjects did. The question arises whether it is fair to use the best human performance as a standard of comparison, and whether this creates a false impression based on the performance of one or two exceptionally skilled individuals. The results dispel this concern, since they indicate that subjects were by and large equally gifted in generating good solutions. Generally, subjects' performances on one problem provided no indication of how they would perform on the next. A measure of this is the average rank-order correlation of performances across different problems, which at .06 was not significantly different from zero. Thus, there was little or no tendency for better or poorer performances to be associated with particular individuals. This is consistent with the previous finding, that performances were tightly clustered in the upper tail of the distribution of problem solutions. With less than three percentile points separating the best and worst of 315 solutions, finding systematic individual differences was extremely unlikely.

It was hypothesized that the complexity of TSPs would be determined by the number of internal points rather than the total number of points. This raised the question of how best to measure problem complexity from the present results. Using the response times for individual problems was ruled out, since the group setting used for data col-

lection prevented measuring solution times for each individual problem. Following the principle that good patterns have few alternatives (Garner, 1970; Pomerantz, 1981), two measures of difficulty were used. The first was the number of different solution paths generated by subjects, which correlated highly with the number of interior points. However, this measure does not distinguish degrees of difference between solutions. A second measure, which does, was based on the response uncertainties associated with all possible connections between pairs of points. For the six basic problems, response uncertainty increased significantly with the number of interior points ( $r = .93$ ), indicating that when the total number of points is held constant, the complexity of a TSP varies directly with the number of interior points.

Also consistent with the hypothesis was the fact that in almost every case subjects connected boundary points in order of adjacency. Of the 315 problem solutions, 310 (98%) were consistent with this principle. The results indicated that almost no uncertainty existed for subjects regarding the order of connection of boundary points. In addition, subjects showed a preference for connecting adjacent boundary points together, and their solutions showed significantly fewer indentations than would be expected.

Nevertheless, consideration of the number of possible indentations for each problem raises another hypothesis regarding problem complexity that could operate as an alternative or as a supplement to the present hypothesis, and one that cannot be ruled out here. In the present problems, the number of possible indentations in a solution ranged from 1 to 5. In the problem with only 1 interior point, only one indentation was possible, whereas in the two problems with 5 interior points the number of indentations in a solution could range from 1 to 5. It is possible that the greater this range, the greater the potential uncertainty associated with how to connect interior points to the boundary. That is, the complexity of TSPs may depend on the maximum number of indentations.

In the present experiment, the maximum numbers of indentations for the problems with 1, 2, 3, 4, 5, and 6 interior points were 1, 2, 3, 4, 5, and 4, respectively. As can be seen, the two variables were highly correlated, and the dependent variable, response uncertainty, correlated almost as highly with the maximum number of indentations ( $r = .85$ ) as it did with the number of interior points ( $r = .93$ ). A second experiment was conducted to try to distinguish the effects of these independent variables. The experiment also provided an opportunity to measure performance in TSPs with a larger number of nodes.

## EXPERIMENT 2

### Method

#### Subjects

The 29 subjects were students at Loughborough University of Technology.

#### Materials

Seven TSP problems with 20 points were generated. The numbers of interior points for the problem were 4, 6, 8, 10, 12, 14, and 16. The corresponding maximum numbers of indentations were 4,

6, 8, 10, 8, 6, and 4. The problems were created in a way similar to that in Experiment 1, except that some constraints were introduced to prevent interior points from falling too closely together. As before, the coordinates of the points on the boundary of the hull were initially determined from the vertices of the regular polygon with that number of sides, using polar coordinates with a radius of 80 mm. Irregularity was again introduced by adding a value selected randomly from the range of +5° to -5° to each angle. Polar coordinates for the interior points were generated randomly with the constraint that they fell not further than 50 mm from and not closer than 10 mm to the center of the original regular polygon. Once 20 points had been generated for each problem, the polar coordinates were converted to rectangular coordinates centered on the coordinates (140, 90) and rounded to integer values. Each problem was drawn on a separate 11 × 8.5 in. sheet. Figure 4 shows the problem with 10 interior points with solutions showing the minimum and maximum number of indentations.

**Design**

The design was similar to that of Experiment 1, with each subject tested on each problem. Two different orders of presentation were used, with the problems presented in the order of 14, 12, 4, 10, 16, 8, and 6 interior points or in reverse order.

**Procedure**

The procedure was the same as in Experiment 1.

**Results**

Nine subjects left one or more of the problems incomplete, and these cases were excluded from analysis, leaving a total of 20 cases.

**Quality of Solutions**

**Standard score approach.** A random sample of 100 solutions was generated for each problem, and the sample means and standard deviations were used to convert subjects' path lengths to standard scores. The main results appear in Table 4, which shows, for each problem, the minimum, mean, and maximum path lengths observed over the 20 subjects. Converting the standard scores to percentiles, all subjects' path lengths for each of the problems were beyond the 99.99th percentile.

**Percentage over optimal approach.** The TRAVEL package was again used to generate solutions to the problems. For five problems, a proven optimal solution was found. For the problems with 4 and 6 interior points, the best solutions found were 2.4% and 1.2% above the

**Table 4**  
Minimum, Mean, and Maximum Path Lengths Produced by Subjects in Experiment 2, and the Corresponding z Values

Number of Interior Points	Subjects' Path Lengths					
	Minimum		Mean		Maximum	
	Score	z	Score	z	Score	z
4	707.49	8.26	724.27	8.15	816.03	7.53
6	703.89	7.16	746.32	7.16	846.50	6.50
8	737.61	7.34	762.40	7.34	830.16	6.82
10	708.93	7.60	721.81	7.60	745.03	7.39
12	692.60	7.79	719.70	7.79	828.50	6.57
14	675.80	7.68	727.36	7.68	830.68	6.42
16	599.72	6.23	644.14	6.23	734.54	5.18

lower bound of an optimal solution, respectively. Therefore, the best solutions found in these cases may or may not be the optimal. For these two cases, the lower bound was used as an estimate of the optimal solution.

The three heuristic algorithms described previously were used with each of the problems, and the path lengths produced by the subjects and the heuristics were converted to a percentage over the optimal. The main results are shown in Table 5. For each problem, the table shows the percentage over the optimal solution of the best solution produced by the subjects and each of the three heuristics. In addition, the table shows the percentage over the optimal for the subjects' mean and the Nearest Neighbor mean (the average of the path lengths produced using each possible starting point).

**Problem Complexity**

The numbers of different solutions generated by subjects were 17, 16, 17, 13, 16, 14, and 18 for the problems with 4, 6, 8, 10, 12, 14, and 16 interior points, respectively. The number of different solutions did not correlate significantly with either the number of interior points or the maximum number of indentations. These results may reflect the problem with using the variety of solutions as a measure of complexity, in that many of the subjects' different solutions were highly similar in both path length and overall configuration.

With response uncertainty as the measure of complexity, a pattern of results emerged that was more consistent with those of Experiment 1. The results are shown

**Table 5**  
Minimum Observed Path Length and Percentage Above the Minimum for Human and Heuristic Solutions, Experiment 2

Number of Interior Points	Optimal Solution	Percentage Above Optimal					
		Subject Minimum	NN Minimum	LA	CHCI	Subject Mean	NN Mean
4	703.81	3.0	4.1	3.1	2.5	5.4	8.4
6	703.89	1.2	1.2	20.1	5.3	7.3	8.8
8	725.31	1.7	3.3	13.7	5.8	5.2	9.4
10	698.83	1.4	0	6.0	4.8	3.3	9.5
12	688.33	0.6	0	23.4	3.6	4.6	7.1
14	663.61	1.9	1.5	10.0	4.4	9.6	15.4
16	593.81	1.0	15.0	0.8	0.05	8.5	21.7

Note—NN, Nearest Neighbor; LA, Largest Interior Angle; CHCI, Convex Hull, with cheapest insertion criterion.



in Figure 3 (open squares), which plots response uncertainty against the number of interior points for each problem. The product-moment correlation between uncertainty and the number of interior points was  $r = .79$  ( $df = 5$ ), significant beyond the .05 level (two-tail test). The corresponding correlation between complexity and the maximum number of indentations per problem was not significant ( $r = -.56$ ) and in fact in the direction opposite to what would be expected if the number of indentations were associated with problem complexity.

### Other Findings

**Individual differences.** Subjects' solutions for each problem were ranked, and the average Spearman rank-order correlation between each pair of rankings was obtained. The result was an average Spearman coefficient of .02 ( $n = 20$ ), indicating again no correlation in subjects' performances across different problems.

**Order of boundary points.** As with the 10-point problems, the large majority of subjects' solutions connected the boundary points in order of adjacency, and only one of the 140 solutions failed to adhere to this principle.

**Crossed arcs.** Additional crossed arcs occurred in five solutions, four of them committed by 1 subject.

**Number of indentations.** Again, subjects produced solutions with fewer indentations than would be expected if they had no preference for connecting adjacent boundary points. The average number of indentations for subjects' solutions was 1.98, less than half the expected mean of 4.11. Seventy-two percent of the solutions exhibited 2 or fewer indentations, compared with an expected frequency of 30%. The difference between the obtained and expected distributions was significant ( $D_{MAX} = .42, p < .01$ ).

### Discussion

The experiment was designed to obtain further information on human performance on TSPs, and to test whether problem complexity is predicted better by the number of interior points or by the maximum number of indentations.

The maximum number of indentations failed to correlate significantly with response uncertainty, while there was again a positive correlation between the number of interior points and complexity. However, the correlation was weaker than in Experiment 1. From the nature of subjects' solutions, it appeared that other factors might have come into play. In particular, the proximity between interior points may have played a greater part in determining connections than in the previous experiment, and it may be that the larger number of interior points in the present case provided more basis for figural factors of this kind to operate. A post hoc regression analysis was conducted in which the average distance between interior points was added as a second independent variable to the number of interior points. This improved the prediction of response uncertainty. The resulting multiple correlation coefficient was .89, significant beyond the .05 level. Taking into account the proximity of interior points as well as their number therefore increased the

variance accounted for by about 17%. While the increase itself was not significant, the results suggest that it may be possible to identify factors in addition to the number of interior points that affect the complexity of TSPs.

With respect to the quality of human solutions, the results corroborated those of the previous experiment. Comparison with a random sample indicated again that subjects' solutions formed a tight cluster in the upper tail of the distribution, with all subjects' solutions falling beyond the 99.99th percentile. The best human solutions again compared favorably with the heuristics. Averaged over the seven problems, the best human solutions were 1.5% above the optimum. The comparable figures for the heuristics were 3.6%, 4.1%, and 11.0% above, for the Nearest Neighbor, CHCI, and Largest Angle procedures, respectively. Once again, the high quality of subject performance could not be attributed to one or two especially skilled individuals. The average rank-order correlation of performances across different problems was virtually zero and indicated no tendency for better or poorer performances to be associated with particular individuals.

### GENERAL DISCUSSION

The experiments were designed to test the hypothesis that the complexity of TSPs for human subjects is a function of the number of interior points rather than of the total number of points. With response uncertainty as a measure of complexity, both experiments supported the hypothesis. The results of the first experiment were also consistent with an alternative hypothesis—that complexity was determined by the maximum number of indentations that a problem permitted—but the second experiment failed to support this alternative, while continuing to support the original hypothesis. However, the correlation between number of interior points and complexity was weaker in the second experiment, where interior points ranged from 4 to 16, than in the first, where they ranged from 1 to 6. There was some indication that increased proximity of interior points in the second experiment may have operated as an additional factor in reducing uncertainty. Nevertheless, the number of interior points appears to be a relatively powerful determinant of problem complexity. The correlation between interior points and response uncertainty across the data from both experiments combined was  $r = .90$  ( $df = 11, p < .01$ ).

The experiments were also attempts to obtain information on the quality of subjects' solutions. No completely satisfactory single measure of quality was available, and multiple approaches were consequently adopted. In the first of these approaches, subjects' path lengths were converted to standard scores, with population parameters estimated from random samples of solutions. In the second, a standard used in operations research was adopted, expressing the obtained path lengths as a percentage above the optimal solution. In the third approach, subjects' solutions were compared with those of heuristic algorithms developed in the field of operations research. All three approaches supported the conclusion that the subjects'

solutions were of a high quality. The standard score approach indicated that solutions were clustered in the upper tail of the distribution, on the average exceeding the 99.9th percentile. Compared with the optimal solutions, subjects' minimum path lengths exceeded them on the average by three-quarters of a percent. As a point of reference, to reliably find solutions within 2% or 3% using computational approaches requires fairly complex algorithms performing in the order of  $n^3$  computations (Golden et al., 1980). Subjects' minimum path lengths were substantially better than this. They were also substantially better than the heuristics tested here, whose best path lengths were 2.5%, 2.5%, and 7.6% above the optimum solutions, for the CHCI, Nearest Neighbor, and Largest Angle procedures, respectively.

It is an open question whether or not subjects can continue to produce good solutions as problem complexity increases. However, unpublished results reported by Lee (1985) appear to extend the range of the present findings. Lee's report provides coordinates for 14 randomly generated problems ranging from 10 to 60 nodes, together with the best solution for each problem obtained from an informal study involving 50 subjects. There were four problems each of 10, 20, and 40 nodes, and two 60-node problems. The number of interior points for the problems ranged from 4 to 50. Unfortunately, one of Lee's results appears likely to be an error. For 1 of the 10-node problems he reports that the best result for both human subjects and the CHCI heuristic was 16% above the optimum. This figure is extremely high compared with his other results and the results found here. Reapplying the CHCI heuristic to his coordinates indicated that it does in fact find the optimum solution, so apparently his figure is an error. If this suspect result is excluded, his results across the remaining 13 problems are highly consistent with those reported here. A reanalysis of his problems indicated that the best human solutions were better than the three heuristics up to and including the 60-node problems. The average percent above the lower bound of the optimal solution was 1.5 for the best human solutions, compared with 2.7, 10.4, and 9.1 for the CHCI, Nearest Neighbor, and Largest Angle heuristics, respectively. In addition, Lee's results showed a fairly high correlation between the number of interior points and the percentage over optimal for the human solutions across the 13 problems ( $r = .84$ ,  $df = 11$ ,  $p < .01$ ). In other words, percentage above the optimal appears to have correlated with problem complexity here. This may be because there were more than a single problem for each level of the independent variable, number of interior points, or because a much wider range of interior points was used (4–50, compared with 1–16 in our experiments).

To return to the present findings—subjects' performances showed a high degree of consistency in a number of respects. Almost invariably, subjects connected boundary points in order of adjacency and, equally invariably, produced no crossed arcs. They also generated solutions with relatively few indentations. These consistent elements, together with the fact that proficiency ap-

pears to be equally distributed across individuals, suggest that solutions are generated by a homogenous process or set of processes. However, none of the heuristics appeared able to provide an adequate model of this process. From what we observed, people complete the problem in a sequential way, in which the terminal node of one connection becomes the origin node of the next. Of the three heuristics that we used, only the Nearest Neighbor approach operates in this way. This approach is similar to human performance in another respect too, in that it can produce a range of solutions to a single problem depending on the starting point selected. The other two heuristics typically produce a single solution. However, the output produced by the Nearest Neighbor heuristic was not particularly close to the human results, and in this respect the procedure does not appear able to provide an accurate model of performance.

The spatial mode of presenting the TSP problem to human subjects appears to be critical to performance, and providing the equivalent information in a nonspatial form, as a set of coordinates, for example, or as a matrix of interpoint distances, is not likely to yield the same quality of solutions, if any solution at all. Experimental evidence for this assertion is provided by Polivanova (1974), who investigated human subjects' performance on route-finding tasks containing between 4 and 10 points. She found that performance with a spatial representation was superior to performance with a tabulated numerical form of the same task. TSPs may be easy for people because people do not need to form an internal representation of the task but simply use the display to prompt potential solutions. The spatial form of presentation of the TSPs may have allowed a form of display-based problem solving (Larkin, 1989), where the current state of the task can be read off from the display and used to decide the next subgoal. However, producing solutions to TSPs through display-based problem-solving implies a local processing strategy, whereby subjects judge the next node to be connected on the basis of the spatial properties of adjacent nodes. This type of local processing model is similar to the Nearest Neighbor approach, and as such, it is unlikely to provide a good account of observed performance. While human solutions are clearly *executed* in a serial fashion, we believe that they are unlikely to be *generated* through the serial application of a display-based problem-solving process. The evidence presented here indicates that human subjects reach solutions based upon the perception of global spatial properties of a TSP, and in particular, of the boundary of the convex hull. Recently, Ormerod and Chronicle (1995) reported evidence supportive of this view, from a recognition task in which subjects were found to reliably distinguish between 100% and 90% optimal solutions for 10-node problems in times of less than 2 sec. The ability to differentiate between such close levels of optimality in such short response times suggests some form of rapid overall processing. There is some evidence that adding contextual information may reduce the quality of subjects' solutions. Buckmaster (1992) reported a pilot study in which

subjects saw points that were superimposed on a map of the United States and identified as U.S. cities. The added geographical context appeared to reduce the quality of subjects' solutions to approximately the level of the Nearest Neighbor heuristic. It appears that the addition of thematic information may have encouraged subjects to adopt a local processing strategy.

The present results provided evidence that subjects follow a convex hull approach in solving TSPs, but without explaining why. One possibility is that the boundary of the convex hull presents itself as the best figure, perhaps providing a two-dimensional illustration of the law of *Prägnanz* in the same way that soap bubbles illustrate it in three (Attneave, 1982). The task of the TSP may happen to parallel what it is natural for the perceptual system to do in any case when presented with an array of dots. For example, Pomerantz reports an experiment in which subjects were presented with dot patterns and were asked to connect the dots in a manner that illustrated how they perceived the patterns (Pomerantz, 1981). He reported that subjects frequently connected the dots by using the shortest possible paths. In other words, the "natural" tendency of Pomerantz's subjects appears to conform closely to the task requirements of the TSP. If this were the case, the factors that make TSPs relatively simple for human subjects may be exactly the same as those which make other problems, such as the 9-dot problem, extremely difficult. If the boundary of the convex hull dominated perception in the 9-dot problem, a subject would presumably find it very difficult to extend lines beyond the hull, as required for successful solution. This is consistent with some explanations that have been offered for the problem's difficulty (Maier, 1930; Scheerer, 1963), though not with all (Lung & Dominowski, 1985; Weisberg & Alba, 1981).

Another possibility is that the apparent influence of the boundary on solutions arises because the problem "invites" such an approach. The overall form of the prob-

lem is essentially circular, and this may have suggested types of solutions that best preserved this overall form. Two versions of this type of explanation can be distinguished, one specific to the present type of stimuli, one general to any TSP task. With regard to the specific version, the points in the problems used here were fairly widely and evenly distributed in the plane and fell within a more or less circular area. This may implicitly have prompted subjects to pay attention to the boundary of the convex hull. That is, the shape of the problem may have suggested the shape of the solution. If this were the case, then changing the overall form of problems should change the nature of solutions. If, for example, the points were enclosable by a more extended, linear form, then, perhaps, other figural factors such as proximity or good continuation might determine subjects' paths. In Figure 5, for example, the question arises of whether subjects starting at the leftmost point would follow the curve and thereby produce the nonoptimal solution on the top, with crossed arcs, or would they continue to be influenced by the boundary of the hull, and find the optimal solution, shown below?

A more general sense in which the TSP is "circular" arises from the task's requirement to complete a circuit, which constrains solutions to the form of a topological ring. It may be that this built-in "circularity" of the task gives the boundary of the convex hull a salience it would not otherwise have. This could be tested by removing the requirement to complete a circuit. For example, the goal could be changed to that of connecting the points in the shortest possible path, starting at the point farthest to the left of the display and ending with the point farthest to the right. This "linear" version of the problem retains all of the local characteristics of the original TSP but alters a global characteristic in a potentially significant way. Figure 6 illustrates three sample solutions to a problem of this type. The solution at the top has only two arcs which connect adjacent boundary points. The lower two

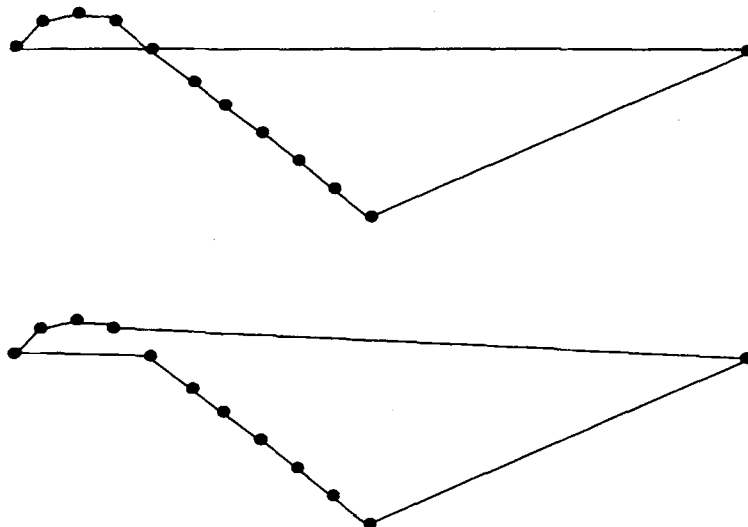


Figure 5. Nonoptimal and optimal solutions to more linearly extended Traveling Salesman Problem.

each have four arcs connecting adjacent boundary points. If the convex hull retains its influence on perception of the problem, then, arguably, subjects should produce solutions more like the lower two than like the top one. If, on the other hand, subjects' solutions conform more to the top one, it would suggest that the salience of the hull in human solutions to the TSP arises from the problem's requirement to complete a circuit.

The present results are essentially exploratory and are subject to some limitations. The problems were drawn from a relatively narrow range in terms of the number of interior points, with only one example at each level. This raises questions about the representativeness of the problems used and limits the generality of the findings. However, this is mitigated to some extent by the inclusion of the benchmark Dantzig problem, which, unlike our quasi-random problems, has a highly structured layout designed to provide a critical test of heuristics. Subjects' solutions to this special problem were similar to those obtained for the other problems, both in quality and in the relationship between number of interior points and problem complexity. A second factor which suggests that the present findings may not be specific to the sample problems used is their consistency with those of Lee (1985) based on a much wider range of randomly generated problems. The latter results indicated that the best human performances still exceeded those of the heuristics up to problems with 60 nodes and 50 interior points.

Only further empirical study will identify the range and types of problems for which human performance continues to be of such a high quality. However, it may be that conditions under which human performance breaks down will be more informative about the mechanisms involved than the conditions under which it remains superior. This need not necessarily require "large" problems. The Dantzig problem has few nodes, but it confounds many heuristic algorithms, and there may similarly be "small" problems that present difficulties for human subjects. Research is currently in progress in which the spatial distribution of internal nodes is systematically varied in the degree of clustering and distance from the boundary.

In conclusion, the present study revealed that human subjects were able to produce very good solutions to an intransigent mathematical problem. The high degree of consistency between individuals argues that a common process or set of processes underlies the ability. A number of factors indicate that the process is perceptually based. It appears not to depend on any overt measurement or calculation. Solutions were produced relatively quickly, requiring roughly 2 min per problem. By way of comparison, computational procedures known to produce solutions of a similar standard would have to perform in the order of 8,000 calculations to solve one 20-node problem. The spatial mode of presenting the problem appears to be critical. It may be that the relative ease with

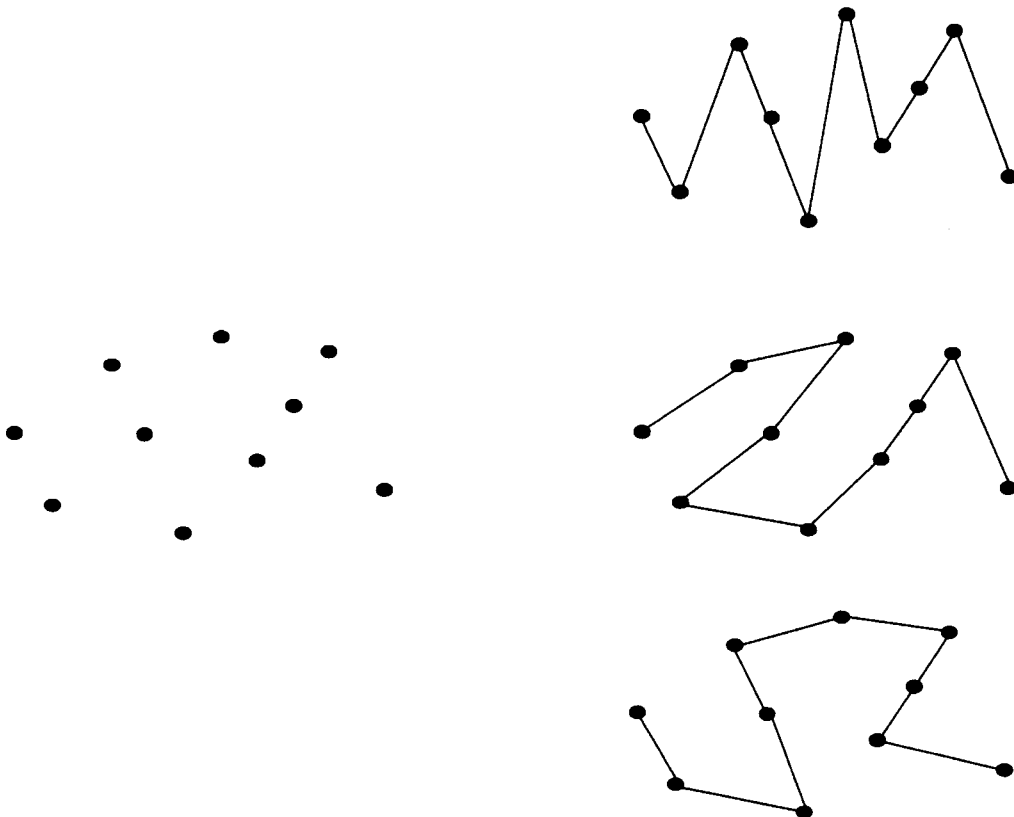


Figure 6. A "linear" Traveling Salesman Problem and three possible solutions.

which subjects produce high-quality solutions arises because the task requirements of the TSP happen to conform to natural tendencies of the perceptual system. However, a theoretical account of human performance will require further research.

REFERENCES

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

BARTON, G. E., JR., BERWICK, R. C., & RISTAD, E. S. (1987). *Computational complexity and natural language*. Cambridge, MA: MIT Press.

BOYD, S. C., PULLEYBLANK, W. R., & CORNUÉJOLS, G. (1987). TRAVEL—an interactive travelling salesman problem package for the IBM personal computer. *Operations Research Letters*, **6**, 141-143.

BUCKMASTER, J. K. (1992). *Evaluating human performance on the travelling salesman problem using a Hypercard based system*. Unpublished master's thesis, Loughborough University of Technology.

DANTZIG, G. B., FULKERSON, D. R., & JOHNSON, S. M. (1959). On a linear-programming, combinatorial approach to the travelling-salesman problem. *Operations Research*, **7**, 58-66.

FLOOD, M. M. (1956). The travelling salesman problem. *Operations Research*, **4**, 61-75.

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

GOLDEN, B., BODIN, L., DOYLE, T., & STEWART, W. (1980). Approximate travelling salesman algorithms. *Operations Research*, **28**, 694-711.

KROLAK, P., FELTS, W., & MARBLE, G. (1971). A man-machine approach toward solving the travelling salesman problem. *Communications of the ACM*, **14**, 327-334.

LARKIN, J. H. (1989). Display-based problem solving. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing* (pp. 319-341). Hillsdale, NJ: Erlbaum.

LEE, R. K. L. (1985). *A heuristic approach to the travelling salesman problem* [Unpublished Management Report]. University of Victoria, School of Public Administration.

LUNG, C.-T., & DOMINOWSKI, R. L. (1985). Effects of strategy instructions and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 804-811.

MAIER, N. R. F. (1930). Reasoning in humans: I. On direction. *Journal of Comparative Psychology*, **100**, 115-143.

MICHIE, D., FLEMING, J. G., & OLDFIELD, J. V. (1968). A comparison of heuristic, interactive and unaided methods of solving a shortest-route problem. In D. Michie (Ed.), *Machine intelligence* (pp. 245-255). New York: Elsevier.

NORBACK, J. P., & LOVE, R. F. (1977). Geometric approaches to solving the travelling salesman problem. *Management Science*, **23**, 1208-1223.

ORMEROD, T. C., & CHRONICLE, E. (1995, September). *Rapid solution to perceptually presented intransigent problems: The case of the travelling salesman*. Paper presented at the British Psychological Society Cognitive Section Conference, Bristol.

POLIVANOVA, N. I. (1974). On some functional and structural features of the visual-intuitive components of a problem-solving process. *Voprosy Psikhologii* [Questions of Psychology], **4**, 41-51.

POMERANTZ, J. R. (1981). Perceptual organization and information processing. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 141-180). Hillsdale, NJ: Erlbaum.

SANGALLI, A. (1992, December 12). Why sales reps pose a hard problem. *New Scientist*, pp. 24-28.

SCHEEERER, M. (1963). Problem solving. *Scientific American*, **208**(4), 118-128.

WEISBERG, R. W., & ALBA, J. W. (1981). An examination of the alleged

role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General*, **110**, 169-192.

WILF, H. S. (1986). *Algorithms and complexity*. Englewood Cliffs, NJ: Prentice-Hall.

APPENDIX

Interior Points	x-, y- Coordinates
10-Node Problems	
6	196.83, 146.3; 81.42, 144.49; 85.83, 31.13; 198.74, 35.69; 114.25, 90.37; 178.62, 79.78; 135.33, 99.61; 120.41, 69.85; 109.31, 91.33; 132.41, 69.11
5	215.06, 117.67; 134.03, 169.78; 62.38, 109.39; 89.26, 28.15; 192.35, 29.51; 128.76, 55.87; 139.44, 90.92; 142.22, 71.20; 152.24, 62.73; 168.57, 103.29
4	219.71, 83.18; 177.94, 160.43; 97.40, 157.72; 60.06, 86.99; 104.47, 18.32; 175.66, 18.39; 155.34, 58.40; 150.41, 126.27; 146.55, 87.41; 127.32, 93.50
3	206.66, 134.24; 137.58, 169.96; 77.94, 140.48; 61.22, 76.07; 111.63, 15.20; 170.66, 16.11; 217.29, 69.34; 168.6, 98.57; 162.65, 81.02; 160.42, 107.24
2	219.95, 92.83; 194.79, 148.29; 143.04, 169.94; 80.85, 143.86; 60.00, 90.03; 79.34, 37.85; 137.32, 10.04; 194.27, 31.23; 171.90, 108.66; 103.42, 93.35
1	219.96, 92.44; 197.63, 145.49; 147.71, 169.63; 96.68, 157.26; 65.90, 120.16; 65.18, 61.69; 105.65, 17.75; 157.03, 11.83; 197.15, 34.02; 177.33, 89.20
Dantzig	220, 170; 60, 170; 60, 10; 220, 10; 220, 90; 180, 130 100, 50; 140, 90; 180, 90; 200, 90
20-Node Problems	
16	200, 142; 80, 141; 84, 34; 194, 32; 176, 94; 157, 100; 168, 124; 143, 119; 127, 123; 125, 120; 109, 118; 109, 101; 102, 85; 122, 82; 129, 72; 129, 59; 153, 44; 147, 73; 158, 80; 173, 83
14	210, 128; 138, 169; 73, 132; 69, 55; 145, 11; 209, 51; 157, 94; 176, 121; 153, 137; 138, 103; 135, 105; 132, 98; 117, 94; 112, 80; 116, 63; 129, 43; 145, 55; 158, 46; 155, 76; 178, 80
12	213, 122; 170, 164; 112, 164; 66, 120; 64, 66; 110, 16; 167, 15; 212, 56; 152, 93; 156, 115; 145, 111; 139, 106; 124, 109; 121, 91; 126, 87; 131, 82; 130, 45; 149, 67; 149, 76; 185, 82
10	214, 117; 191, 151; 136, 169; 94, 154; 64, 113; 65, 63; 94, 26; 146, 11; 185, 25; 216, 67; 159, 94; 153, 106; 136, 129; 123, 103; 108, 97; 114, 85; 129, 81; 132, 46; 157, 70; 156, 84
8	215, 115; 192, 149; 164, 166; 122, 167; 82, 144; 64, 112; 62, 76; 84, 34; 114, 15; 166, 15; 195, 33; 218, 74; 173, 111; 148, 105; 122, 113; 99, 92; 130, 88; 125, 59; 147, 56; 167, 82
6	217, 109; 199, 143; 168, 164; 146, 169; 109, 163; 76, 137; 62, 105; 62, 77; 77, 42; 107, 18; 139, 11; 172, 17; 198, 35; 217, 69; 182, 114; 142, 111; 112, 97; 114, 82; 147, 47; 150, 87
4	217, 111; 210, 128; 187, 154; 148, 169; 131, 169; 97, 157; 76, 137; 62, 105; 61, 80; 74, 46; 95, 25; 125, 12; 155, 12; 178, 20; 209, 50; 219, 74; 182, 98; 127, 111; 133, 42; 161, 63

(Manuscript received February 14, 1995; revision accepted for publication September 14, 1995.)