Title
Use of clustering in human solutions of the traveling salesperson problem

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Marupudi, Vijay
Harsch, Rina
Rao, V. N. Vimal
et al.

Publication Date
2022

Peer reviewed

Use of clustering in human solutions of the traveling salesperson problem

Vijay Marupudi (vijaymarupudi@gatech.edu)
School of Interactive Computing
Atlanta, GA 30308 USA

Rina Miyata Harsch (harsch@umn.edu)
Department of Educational Psychology,
Minneapolis, MN 55455 USA

Jimin Park (park1481@umn.edu)
Department of Educational Psychology,
Minneapolis, MN 55455 USA

V. N. Vimal Rao (rao00013@umn.edu)
Department of Educational Psychology,
Minneapolis, MN 55455 USA

Jeffrey Kramer Bye (jbye@umn.edu)
Department of Educational Psychology,
Minneapolis, MN 55455 USA

Sashank Varma (varma@gatech.edu)
School of Interactive Computing
Atlanta, GA 30308 USA

Abstract

The traveling salesperson problem (TSP) is an NP-Hard problem that computers find difficult to solve. Humans are surprisingly good at solving the TSP, with solutions within 10% of optimal for problems with up to 100 points, constructed in time linear with the number of points. We propose that humans solve the TSP by initially clustering the points and then connecting them first within and then between clusters. In this study, 67 participants first clustered 40 stimuli and then solved them as TSPs. Strikingly, participants' TSP solutions *perfectly* followed their clusters for 52% of the stimuli. Further, participants' TSP solutions were more congruent with their clusters for stimuli with statistically higher levels of clustered structure. This provides strong evidence for the clustering proposal. Random TSP solutions, however, showed no such congruence to cluster structure. These findings suggest that clustering might be a fundamental ability for human reasoning about graph-theoretic algorithmic problems.

Keywords: traveling salesperson problem; clustering; problem solving; computational complexity; computational thinking

Introduction

The traveling salesperson problem (TSP), one of the most well-studied optimization problems of computer science, is an NP-hard problem that is difficult for computers to solve. Solving a TSP for a given a set of points involves finding the shortest tour that visits each point once and returns to the starting point. The difficulty of the TSP increases exponentially with the number of points, leading to a combinatorial explosion of time and space costs when solving large problems. Efficient solutions to TSPs have wide applicability for a variety of fields such as package delivery, transportation, DNA sequencing, and the development of semiconductors, making it an important problem to understand.

Humans routinely face and (approximately) solve problems that are difficult for computers, but humans' processes are not yet well understood. Therefore, understanding how humans solve the TSP could provide more general insights into the heuristics they use and the constraints they place on hard problems to make them more tractable, leading to a richer computational theory of human problem solving (Van Rooij et al., 2012).

Humans have been found to be surprisingly good at solving the TSP, with fast solution times that increase linearly with the number of points (Graham et al., 2000). People can solve TSPs of up to 100 points while staying within 10% of the optimal solution (Dry et al., 2006). A study by van Rooij et al. (2006) showed that even 7-year-old children can solve TSPs with decent performance, and the quality of TSP solutions increased with participants' age to adulthood. Together, these findings suggest that perceptual processing may be one capacity used to solve TSPs, and increased analytical processing in adults improves TSP performance. Macgregor et al. (2000) presented a computational model of humans' TSP performance suggesting that their problem solving is guided by the convex hull (the "border" of a problem, i.e., the smallest shape that contains all of its points). They proposed that people generally select points on the convex hull, while picking up more interior points along the way. The model performed better at predicting the length of human TSP solutions compared to the naive nearest neighbor algorithm, where the closest available point is chosen next.

However, there are some limitations to this approach. Problems with more points on the convex hull should be easier according to this model, but humans sometimes judge them as more difficult (Dry & Fontaine, 2014). Additionally, model fit was evaluated based on the difference between tour lengths generated by the model vs. humans, which may be too coarse a metric because qualitatively different solutions for a TSP instance can have similar tour lengths. Perhaps the greatest challenge to this model is that the worst-case time complexity of computing the convex hull of a set of n points is $O(n \log n)$ for problems on the plane, a greater complexity than the linear time humans usually take.

Here, we consider a different proposal: that people solve TSPs by first *clustering* their points. This decomposes a larger problem into a set of smaller problems. Next, they choose an initial cluster and a starting point within, then connect the points in the cluster. This is efficient because clustered points are, by definition, fewer and closer together. When they finish connecting the points in a cluster, they jump to the next cluster and repeat this process. They continue until all points within all clusters are connected, finally finishing at the starting point in the initial cluster. We are not the first to propose that people use clustering to guide the solu-

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

tion of TSPs. Graham et al. (2000) outline their "Pyramid" model, inspired the by the hierarchical structure and parallel processing of the human visual system, which produces tours of lengths similar to those human participants. However, like Macgregor et al. (2000), the specific tours generated by their model do not necessarily resemble human solutions of the same TSPs. Kong and Schunn (2007) offer a model that uses the K-Means clustering algorithm to implement a strategy similar to that of the Graham et al. (2000) model, and which achieves high path correlations with human solutions for the TSP.

A major limitation of clustering-based models of human TSP performance is that human clustering of dot arrays had not been rigorously studied until very recently. This was an important gap, as it was unclear whether humans possess a stable clustering ability which can in turn serve as the basis for the efficient solution of TSP problems. Recent work in our lab has shown that people are highly reliable at clustering dot arrays, suggesting a stable algorithm (Marupudi et al., 2020). Participants were asked to cluster the same stimulus twice at different time points. We found remarkable stability in performance: They tended to include the same points in the same clusters on both occasions. This reliability varied by how statistically clustered (vs. dispersed) the stimulus was: More clustered stimuli were more reliably clustered than more dispersed stimuli. We subsequently found a similar trend in the reliability of human TSP solutions, with more reliable TSP paths for clustered vs. dispersed stimuli (Marupudi et al., 2021). Finding reliability in both clustering and TSP performance is consistent with the use of clustering as a step in solving TSP instances.

The current study builds on prior work from our lab and from other investigations of clustering as foundational to TSP problem solving. Importantly, this is the first study to *directly* tie the cluster performance and the TSP performance of a particular individual on a particular stimulus. Participants saw each stimulus twice, once as a clustering problem and once as a TSP problem. We analyzed whether their TSP solutions followed the contours of their clustering solutions. In addition to this structural alignment, we also looked for evidence of a clustering mechanism in the temporal dynamics of TSP problem solving. Specifically, we predicted that people would take longer to connect points in different clusters than points in the same cluster (after controlling for differences in the lengths of inter-cluster versus intra-cluster connections). Finally, we compared the clustering and TSP solution processes of humans against a novel set of baselines: randomly generated clusters, K-Means generated clusters, and solutions generated by optimal TSP solver programs.

Methods

Participants

Sixty-seven undergraduate students at a large public university in the Midwest U.S. completed the study. Participants were given 1 hour to complete the study (*Median* = 44.7 min.) and were compensated with a $15 gift card. The study was approved by the local IRB.

Design

Both the Clustering and TSP tasks followed a 5 × 2 within-subjects design. The factors were Number of Points (10, 15, 20, 25, 30) and Cluster Structure (clustered, dispersed), varied orthogonally. Four stimuli were generated for each of the ten cells of the design, as described next. For the Clustering task, we recorded each participant's clusters for each stimulus, i.e., the cluster membership of each point. For the TSP task, we recorded each participant's tour for each stimulus, i.e., the sequence of edges, as well as the time to construct each edge. These measurements were the basis for the dependent variables we analyzed, as described below.

Materials

Stimuli were generated randomly using a uniform distribution across a two-dimensional 800 x 500 pixel canvas, and were then filtered for having the desired amount of Cluster Structure. We used the Z-score index of the amount of cluster structure in a stimulus. This includes the variance and edge effect estimates provided by Donnelly (1978), defined as:

$$Z = \frac{\bar{d} - E(d_i)}{\sqrt{Var(\bar{d})}}$$

$\bar{d}$ is the nearest neighbor distance:

$$\bar{d} = \frac{\sum_{i=1}^{N} d_i}{N}$$

$E(d_i)$ is the expected value of the nearest neighbor distance for random patterns where A is the area and B is the perimeter of the Canvas:

$$E(d_i) = 0.5\sqrt{\frac{A}{N}} + \left(0.0514 + \frac{0.041}{\sqrt{N}}\right)\frac{B}{N}$$

Finally, $Var(\bar{d})$ is defined as:

$$Var(\bar{d}) = 0.070\frac{A}{N^2} + 0.037B\sqrt{\frac{A}{N^5}}$$

See Donnelly (1978) and Ripley (1979) for more information on the Z-score index, its definition, and the constants in the formulae above.

We made alterations to the Z-score index to account for margins in the stimuli. These were necessary because the metric was not originally designed for guiding the design of experimental materials. Pilot testing showed that participants ignored the whitespace around the perimeter of a stimulus. This made the raw Z-score index inaccurate when randomly generated points happened to result in sizable margins on the canvas. To control for this, we calculated the Z-score index for stimuli after first removing the whitespace margins.

Then, for each level of the Number of Points factor (i.e., 10, 15, 20, 25, and 30 points), we generated images with the

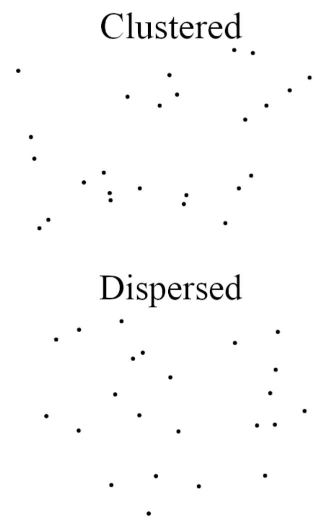

Figure 1: Examples of clustered (Z-score: 1.009) and dispersed (Z score: -2.003) instances.

specified number of points positioned randomly on the two-dimensional 800 x 500 pixel canvas. We then randomly selected four images with corrected Z-score index values in the range 1.00 ± 0.05 as the clustered stimuli, and four images with values in the range -2 $\pm$ 0.05 as the dispersed stimuli. We chose these ranges based on pilot testing, with the goal to select clustered stimuli that did not appear obviously clustered when viewed in isolation, but for which their cluster structure became clearer when they were directly contrasted with the dispersed stimuli (see Figure 1). This resulted in 40 unique stimuli for participants to view in both the Clustering and TSP tasks.

Procedure

Participants first clustered all 40 stimuli, presented in a random order, using a custom plugin implemented in jsPsych (De Leeuw, 2015). For each stimulus, participants were asked to draw an enclosing circle around each set of points that formed a cluster. Points turned blue when participants enclosed them in a cluster. If a point was included in multiple enclosures, it was uniquely assigned to the first cluster. Participants were prohibited from drawing a single cluster around all points, and they could not undo clusters once drawn. A trial ended when all of the points had been enclosed in a cluster. We recorded the cluster membership of each point, the number of clusters drawn, and the timestamp to draw each cluster.

Participants then completed an unrelated math task as a distractor task for 5 minutes (calculating the value of factorial expressions like $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$).

Finally, participants saw the same 40 stimuli again, in a new random order, but this time presented as TSP problems. Half the stimuli were flipped horizontally and vertically from their orientation during the Clustering task. They solved each

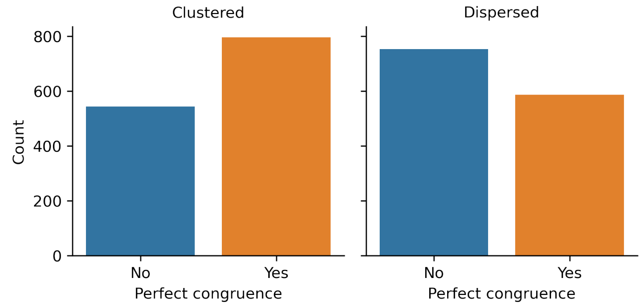

Figure 2: Prevalence of perfectly congruent trials (Cluster deviance = 0) among clustered and dispersed stimuli.

problem using a custom jsPsych plugin by clicking on each point in the tour, in order. Following each click, the point turned blue, and an edge was drawn in blue from the previous point to the clicked point. Thus, the tour was shown to participants as they constructed it. The trial ended after participants selected the last remaining point. Again, they could not undo their selections. We collected timestamps for each clicking of a point, and also recorded their overall tours.

Results

Quantifying Cluster Deviance and Congruence

To quantify how different participants' TSP solutions were from their clusterings of the same stimulus, we created a measure of *cluster deviance*. Our initial metric first calculated the number of cluster transitions, i.e., the number of times a participant's TSP tour crossed from a point in one cluster to another cluster, based on the clusters they individually defined for the same stimulus earlier. By definition, the minimum number of cluster transitions is the number of clusters the participant had drawn, and it occurs only when a participant's TSP tour perfectly respects their own clustering. To anchor this measure at 0, we subtracted off the number of clusters the participant had drawn. Thus, a value of 0 indicates minimal cluster deviance (i.e., maximal congruence), and higher values indicate increasing cluster deviance (i.e., decreasing congruence).

A problem with this metric is that its maximal possible value increases with the number of points (and thus the number of possible deviant transitions), making comparisons across stimuli difficult. We therefore normalized this "raw" metric by dividing by the maximal deviation score for a stimulus. This produced the final cluster deviance metric:

$$\frac{t-c}{n-c}$$

where t is the number of cross-cluster transitions in TSP, c is the number of clusters, and n is the total number of points. This metric ranges from 0 (perfect congruence between a TSP solution and a clustering) to 1 (maximum deviance).

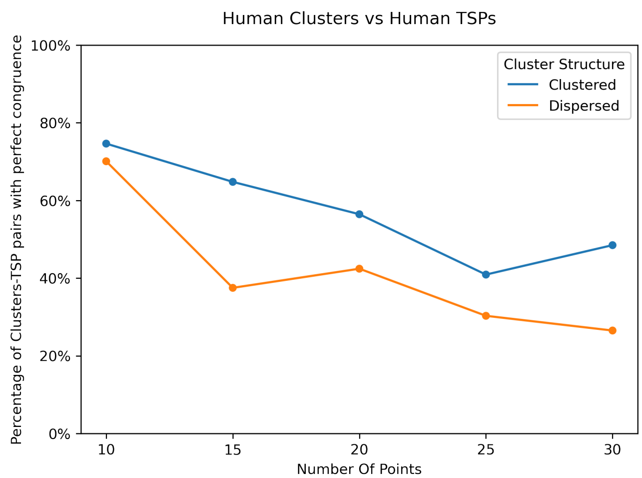

Figure 3: Relationship between Number Of Points and percentage of perfectly congruent Human Clusters-TSP pairs. The pairs maintain surprisingly high levels of perfect congruence for all numbers of points presented to participants, with statistically clustered stimuli showing more cases of perfect congruence.

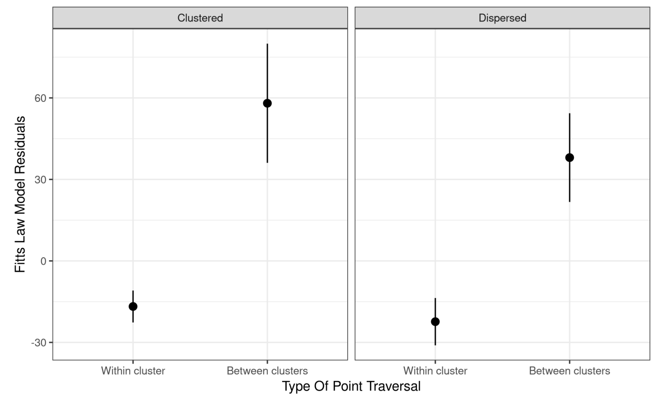

Figure 4: Reaction time per point during participant TSP solutions, controlled for motor ability using Fitts' Law, as a function of whether the point clicked was within or between a participant's clusters.

Correspondence Between TSP Tours and Clustering Solutions

Our major prediction is that people's TSP solutions will follow their clusterings of the same stimuli. We first looked at the overall distribution of cluster deviance scores across all participants and all stimuli plotted separately for the clustered versus dispersed stimuli. Cluster deviance scores were very low. Remarkably, for 52% of all stimuli, the cluster deviance was 0, signaling *perfect* congruence between an individual's TSP solution and clustering. As predicted, congruence was worse for the dispersed stimuli than for the clustered stimuli. For the former, for 75% of the stimuli, the cluster deviance was less than 0.125; for the latter, for 75% of the stimuli, the cluster deviance was less than 0.0714.

To statistically evaluate these results, we counted the number of perfectly congruent TSP-clustering pairs and the number of deviant pairs, separately for the clustered versus dispersed stimuli (Figure 2). We found that clustered stimuli had significantly more perfect pairs compared to dispersed stimuli ($\chi^2(1) = 64.63, p < 0.001$). This is consistent with our proposal that people solve TSP problems by first clustering the points. Since previous work (Marupudi et al., 2021) showed that clustering is less reliable for more dispersed stimuli, there is a greater probability that it produces a different result when applied first during the clustering task and then during the TSP task, resulting in fewer perfect congruence scores. Nevertheless, it should be noted that participants' TSP solutions still maintained a high level of congruency with their clusterings of dispersed stimuli in an absolute sense: 43% of the dispersed stimuli resulted in perfect congruence scores.

To explore the impact of increasing numbers of points, we plotted the number of perfectly congruent TSP-clustering pairs as a function of this variable; see Figure 3. We see that perfect congruence drops sharply for dispersed stimuli between 10 and 15 points. By contrast, perfect congruence for clustered stimuli mostly remains above 50% of trials for the entire range of n explored in the current study.

Temporal Dynamics of TSP Problem Solving

Our proposal is that participants solve the TSP by first clustering the stimulus, connecting the points within the initial cluster, then connecting to a point in the next cluster, and repeating the process. This leads to the prediction that connecting the points within a cluster should be faster than connecting points between clusters, as the participants have already identified the current cluster as their focus. By contrast, connecting to a point in the next cluster should be slower because it requires first choosing the next cluster among the clusters that remain, and then shifting one's focus.

To validate this prediction, we looked at the amount of time spent by participants connecting points within clusters versus between clusters. However, there is a confound between the cognitive proposal above and the physical structure of the stimuli: By definition, points within the same cluster are physically closer to each other than points in different clusters. Therefore, a time difference might reflect not a cognitive process but rather the physics of movement.

To control for this possible confound, we first fit a Fitts' law model for each participant, predicting the time they would take to move from the current point to the next point (i.e., to connect them) as a function of (1) the distance of the movement and (2) the size of the next point. This accounted for the motor movement component of their times. We then collected the residuals and fit a linear mixed effects model predicting them from the cognitive variables of interest: whether a connection was made within or or between clusters, whether the stimulus was clustered or dispersed, and

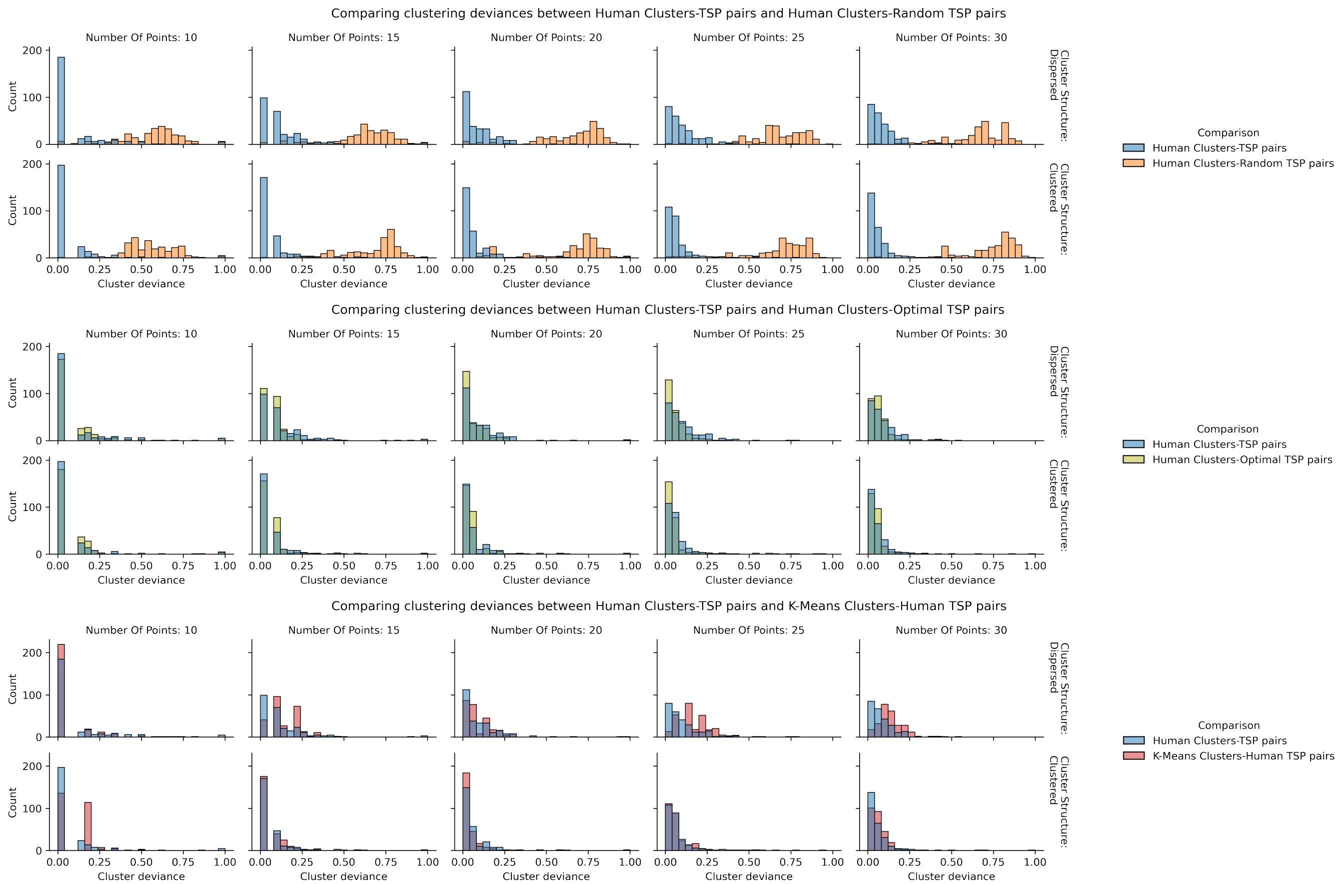

Figure 5: Distributions of cluster deviance across Number Of Points and statistical Cluster Structure. Cluster deviances computed between human cluster-TSP pairs shown in blue in each panel, compared against Human Clusters-Random TSP (top), Human Clusters-Optimal TSP (middle), and K-Means Clusters-Human TSP (bottom). Deviance values close to 0 indicate more TSP solution conformity with clusters. Random TSPs and K-Means clusters compared to human counterparts are more deviant compared to the Human Clusters-Human TSP baseline while optimal TSP solutions are less deviant from human clusters for dispersed stimuli than human TSP solutions themselves.

an interaction between the two. Finally, we added all predictors (including the intercept) as random effects for each participant to account for individual differences. The model provided evidence ($t = 4.010$, $p < 0.0001$) that participants took 50+ ms longer to connect points in different clusters than points within the same cluster; see Figure 4. This is evidence that the cluster structure of TSP problems governs the temporal dynamics of incremental tour generation.

Comparison of Human Performance to Formal Models

We found that 52% of human TSP solutions correspond perfectly with their clusterings (Figure 2). This seems to be a remarkably high percentage, and thus strong support for our clustering proposal — but is it? Comparison to a 'null hypothesis' helps to clarify this finding. We derived one by computing the deviance scores between human clusterings and randomly generated TSP solutions. The first and second rows of Figure 5 show the distribution of these scores separately for the clustered and dispersed stimuli. It is immediately clear that the random TSPs diverged greatly from the human clusterings, with almost no pairs showing perfect congruence. This contrasts strongly with the cluster deviance scores computed over the human clusterings and human TSP solutions.

We then considered another approach to quantify the goodness of human TSP performance. Traditionally, this has been stated as "humans generate tours that are within X% of the minimum length". We can also evaluate their goodness relative to our clustering proposal, by evaluating whether the clusterings people generate are consistent with optimal TSP tours for the same stimuli. To do so, we computed the cluster deviance between a participant's clustering of a given stimulus and the optimal TSP tour as given by the Concorde solver (Applegate et al., 2006). The results are shown in the third and fourth rows of Figure 5. Note that the distribution of cluster deviances, computed separately for the clustered and the dispersed stimuli, is comparable to the human data. Interestingly, we see that the optimal TSP solution tracks the participants' initial clusters for dispersed stimuli more closely than the participants' TSPs for dispersed stimuli themselves.

One hypothesis for this result could be that people recluster the stimulus after connecting the points within each cluster, and the variability in this process might lead to relatively suboptimal TSPs. This would also explain why it took people longer to select points between clusters compared to within clusters. Incremental clustering in this fashion would lack the attention to "global" detail that whole stimulus clustering would entail, and could have led to poorer quality clusters (and TSPs) as a result. This strategy can be adaptive; it prevents people from having to remember their "global" clusters, as they could recalculate clusters quickly when they need them. This result could also indicate the presence of alternate strategies followed by participants when they cluster dispersed stimuli.

We also evaluated human near-optimality from the converse directon by asking whether participants' TSP solutions deviated from clusters not generated by the participants themselves, but by the statistical algorithm K-Means (Lloyd, 1982; Pedregosa et al., 2011). For each participant and each stimulus, we used the number of clusters the participants identified as the input value K to K-Means, generated a clustering, and computed its deviation from the TSP solution the participant generated for the same stimulus. The distributions of these cluster deviances is shown in the fifth and sixth rows of Figure 5 separately for the clustered stimuli and the dispersed stimuli. Here, we see more deviation from statistical optimality. When comparing the K-means clusters with participants' TSP solutions, we find that they are not as congruent with the TSPs as the participants' clusters are. Perfect congruence drops off more sharply for both disperse and clustered stimuli, while human-human Cluster-TSPs maintain high levels of congruence even at 30 points.

Discussion

Previous studies of human solutions of the traveling salesperson problem have largely tried to infer strategies using the optimality of the TSP solutions they generate and the time they take to do so, with the exception of Kong and Schunn (2007). Posited strategies include the convex-hull hypothesis (Macgregor et al., 2000), the avoidance of crossings heuristic (Van Rooij et al., 2003), and the clustering hypotheses (Graham et al., 2000; Kong & Schunn, 2007). In this study, we took a more systematic approach to investigating the clustering hypothesis, showing that the clusters a participant sees in a stimulus strongly structure their TSP solution for the same stimulus.

The current research built on our prior work which showed that participants' clusterings and their TSP solutions are reliable when completed at different time points for the same stimulus (Marupudi et al., 2021; Marupudi et al., 2020). The reliabilities were generally high regardless of the cluster structure of the stimulus, though they were higher for statistically clustered stimuli compared to dispersed stimuli. Increasing the number of points reduced reliability for dispersed stimuli but not for clustered stimuli. These findings provided indirect evidence suggesting that participants might have a stable clustering ability, and may use it to solve TSP problems.

The current study provides more *direct* evidence for this proposal. We found remarkable convergence between participants' clusters and their TSP solutions on the same stimuli. 52% of participants' TSP solutions were perfectly congruent with their clusterings of the same stimuli. This was true even for the dispersed stimuli, with a 43% rate of perfect congruence. In accordance with patterns observed in past research, we found that increasing the number of points results in a decrease in perfect congruence for the dispersed stimuli, although relativly high congruence is maintained for the clustered stimuli. Finding the same patterns in clustering reliablity, TSP reliability, and clustering-TSP congruence suggests use of the same clustering mechanism in all three cases.

Additionally, we compared human-made clusters with randomly generated TSP solutions and with (algorithmically generated) optimal TSP solutions, and we compared human TSP solutions with K-Means-generated clusters. These comparisons led to interesting insights about the optimality of the strategies adopted by participants. Strikingly, optimal TSPs followed participants' clusters more closely than their own TSPs did on dispersed stimuli. On the other hand, K-Means-generated clusters performed poorly at predicting participants' TSP solutions compared to participants' own clusters. This could be because K-Means, as a clustering algorithm, does not generally model human clustering, even after providing the value of K from the participants' clusterings. This could also imply that people's clusterings are well-suited for TSP-like tasks, perhaps because humans routinely use clustering in daily life to plan and solve problems, e.g., when performing multiple errands in a part of town that is new to them.

This study raises important questions about strategy use during TSP problem solving. When participants diverge from their originally drawn clusters, it is unclear if this is due to the unreliabilty of clustering, or whether this is evidence for use of a different, non-clustering strategy. Eye-tracking studies and computational models may provide insight here. Variance in strategy use could also point to individual differences. Vickers et al. (2001) presented evidence of stable individual differences in the quality of TSP solutions that participants provide. Further research can investigate whether these differences arise due to differences in strategy use or due to variability in clustering ability (or both).

Clustering, a form of unsupervised learning, has also been implicated in category learning (Bröker et al., 2022) and language learning (Swingley, 2005). Here, we find evidence that it might be an important ability for computational thinking more generally and for reasoning about graph-theoretic problems, such as the TSP in particular.

Acknowledgements

We thank Kori Terkelsen for coordinating data collection on this study.

References

Applegate, D., Bixby, R., Chvatal, V., & Cook, W. (2006). *Concorde TSP solver*.

Bröker, F., Love, B. C., & Dayan, P. (2022). When unsupervised training benefits category learning. *Cognition*, *221*, 104984. https://doi.org/10.1016/j.cognition.2021.104984

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior research methods*, *47*(1), 1–12.

Donnelly, K. (1978). Simulation to determine the variance and edge-effect of total nearest neighbour distance. *Simulation methods in archeology*.

Dry, M., & Fontaine, E. (2014). Fast and Efficient Discrimination of Traveling Salesperson Problem Stimulus Difficulty. *The Journal of Problem Solving*, *7*(1), 9.

Dry, M., Lee, M., Vickers, D., & Hughes, P. (2006). Human Performance on Visually Presented Traveling Salesperson Problems with Varying Numbers of Nodes. *The Journal of Problem Solving*, *1*(1). https://doi.org/10.7771/1932-6246.1004

Graham, S. M., Joshi, A., & Pizlo, Z. (2000). The traveling salesman problem: A hierarchical model. *Memory & cognition*, *28*(7), 1191–1204.

Kong, X., & Schunn, C. D. (2007). Global vs. local information processing in visual/spatial problem solving: The case of traveling salesman problem. *Cognitive Systems Research*, *8*(3), 192–207.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489

Macgregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2000). A model of human performance on the traveling salesperson problem. *Memory & Cognition*, *28*(7), 1183–1190. https://doi.org/10.3758/BF03211819

Marupudi, V., Harsch, R., Rao, V. N., Bye, J. K., Park, J., & Varma, S. (2021). The role of clustering in the efficient solution of small Traveling Salesperson Problems. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*(43).

Marupudi, V., Rao, V. N. V., Park, J., Harsch, R. M., Bye, J. K., & Varma, S. (2020). Clustering as a precursor to efficient and near-optimal solution of small instances of the Traveling Salesperson Problem (TSP). In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Ripley, B. D. (1979). Tests of 'Randomness' for Spatial Point Patterns. *Journal of the Royal Statistical Society. Series B (Methodological)*, *41*(3), 368–374. https://doi.org/10.1111/j.2517-6161.1979.tb01091.x

Swingley, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive psychology*, *50*(1), 86–132.

Van Rooij, I., Stege, U., & Schactman, A. (2003). Convex hull and tour crossings in the Euclidean traveling salesperson problem: Implications for human performance studies. *Memory & cognition*, *31*(2), 215–220.

Van Rooij, I., Wright, C. D., & Wareham, T. (2012). Intractability and the use of heuristics in psychological explanations. *Synthese*, *187*(2), 471–487.

van Rooij, I., Schactman, A., Kadlec, H., & Stege, U. (2006). Perceptual or Analytical processing? Evidence from children's and adult's performance on the Euclidean traveling salesperson problem. *The Journal of Problem Solving*, *1*(1), 6.

Vickers, D., Butavicius, M., Lee, M., & Medvedev, A. (2001). Human performance on visually presented Traveling Salesman problems. *Psychological Research*, *65*(1), 34–45. https://doi.org/10.1007/s004260000031