In D. Reitter & F. E. Ritter (Eds.), Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016). University Park, PA: Penn State.

Considerations Influencing Human TSP Solutions and Modeling Implications

Brandon S. Perelman (Brandon.S.Perelman.CTR@mail.mil)
U.S. Army Research Laboratory, Human Research & Engineering Directorate
Aberdeen Proving Ground

Shane T. Mueller (ShaneM@MTU.edu)
Michigan Technological University, Department of Cognitive and Learning Sciences,
1400 Townsend Drive Houghton, MI 49931

Abstract

The visual Euclidean Traveling Salesman Problem (TSP) presents participants with nodes, representing cities, and requires that the participant trace the shortest closed route among the cities. Humans solve a similar problem in every day navigation and search tasks. We investigated human TSP solutions for considerations other than solution length. We found a preference for solutions favoring distance-discounted reward and distance to first contact. A hierarchical stochastic model parameterizing solution length, distance-discounted reward, goodness of fit, and plan complexity showed similar effects. The model shows promise for approximating human performance in TSP and other TSP-like naturalistic tasks.

Keywords: TSP; planning; problem solving; visual cognition.

Introduction

The Traveling Salesman Problem (TSP) is a spatial combinatorial optimization problem used in various forms in applied settings, such as operations (e.g., the Vehicle Routing Problem; Dantzig & Ramser, 1959) and engineering (Krolak, Felts, & Marble, 1971), and basic research on spatial cognition and navigation in animals (de Jong, Gereke, Martin, & Fellous, 2011) and humans (Tenbrink & Seifert, 2011). Visual Euclidean TSP requires that the solver plot the shortest path through a 2D metric space containing nodes, representing cities, beginning and ending in the same location. TSP is computationally intractable, with each problem having $(n - 1)! / 2$ solutions. Therefore, brute-force approaches to obtaining optimal, shortest path solutions are too resource-intensive for many applications.

Despite the aforementioned complexity of TSP, human solutions to TSP are typically an order of magnitude shorter (i.e., better) than those produced by many heuristic algorithms (MacGregor & Ormerod, 1996), and are typically no more than 10% longer than the optimal solutions, increasing linearly with problem size (Dry, Lee, Vickers, & Hughes, 2006; MacGregor & Ormerod, 1996; Pizlo et al., 2006). Because human solutions are fast and near-optimal, understanding the mechanism people use to generate them has implications for algorithm development.

Evidence suggests that humans do not exhaustively solve the problem at initial presentation. For example, Kong and Schunn (2007) showed that participants perform the majority of their global information-seeking saccades after beginning to solve the problem. Mueller, Perelman, Tan, and Thanasuan (2015) found very short (~4s) planning times (interval between initial viewing and beginning to solve the problem) that increased linearly with problem size.

These characteristics have prompted the suggestion that humans use a hierarchical approach to problem solving in which a rapidly formed global plan guides the local decisions (e.g., Best & Simon, 2000). Many computational accounts of TSP follow this hierarchical structure, simplifying the problem space by grouping individual cities into clusters (e.g., Pizlo et al., 2006) or designating a global path through the space that starts as a convex hull (MacGregor, Ormerod, & Chronicle, 2000).

This same strategy of following a rapidly produced global plan is likely used in similar tasks. One such task, searching for a target among candidate locations, requires planning a route that optimizes a distance-discounted reward function to minimize the estimated time to find (ETF) that target (see Wiener, Schnee, & Mallot, 2004). This general task is critical in operational domains, such as wilderness search and rescue (Perelman & Mueller, 2013) and military and public safety search operations (Antoniades, Kim, & Sastry, 2003), as well as for sports such as orienteering (Blum et al., 2007). The present study investigates the extent to which a single adaptive mechanism could be used to solve TSP and other TSP-like problems.

Investigations of human behavior in naturalistic TSP-like tasks (e.g., Blum et al., 2007; Perelman & Mueller, 2013; Perelman & Mueller, 2015; Ragni & Wiener, 2012; Tenbrink & Seifert, 2011; Tenbrink & Wiener, 2009) suggest that real-world strategic planning requires considering factors other than path length. Many of these tasks require solvers to prioritize ETF and distance to first contact (DFC), optimizing a function that rewards visiting locations early in the path, versus TSP where rewards are uniform. Wiener et al. (2004) suggest that certain cluster-based strategies should produce this behavior, and Tenbrink and Wiener (2009) found a slight bias (roughly 58% of solutions; 8% more than expected) toward prioritizing larger over smaller clusters early in solutions to a naturalistic TSP-like task. Note that we have termed these alternative solution criteria 'considerations' rather than heuristics as they are not necessarily isolated mechanisms, but components of an underlying mechanism.

If a common mechanism is used to solve TSP and similar tasks, then we should see evidence of these alternative considerations in TSP solutions. To our knowledge, there is no published literature searching for evidence of these considerations in traditional TSP solutions. Evidence of a common mechanism holds implications for algorithm development and our understanding of human visual problem solving.

General Method

The present study consists of analyses of three datasets - two derived from experiments presented here, and one generously donated by other authors (see below).

Experiments (Datasets) 1 and 2

Michigan Technological University students participated in Experiments 1 and 2 (n = 29 and 35, respectively). The goal of Experiment 2 was to replicate the results of Experiment 1 using a blocked design to reduce potential fatigue.

Two participants in Experiment 1 provided incomplete data (final n = 27). Participants completed TSP problems presented using the Psychology Experiment Building Language v. 0.14 (PEBL; Mueller, 2014) TSP. PEBL TSP problems begin in a fixed starting location, with the last segment automatically completed by the software, and route edits are not allowed (see Mueller et al., 2015).

Participants in Experiment 1 completed 5 6-city practice problems, then 15-problem sets presented in random order, each containing 10, 20, or 30 cities for 50 total trials. Participants in Experiment 2 completed their trials in 2 blocks, each containing 5 6-city practice problems, then 5-problem sets each containing 10, 20, and 30 cities presented in random order, for a total of 40 trials between both blocks.

Dataset 3

Data for this analysis were provided by Chronicle, MacGregor, Lee, Ormerod, and Hughes (2008). In that study, 110 University of Hawaii students completed 9 30-city problems by connecting the cities using pen and paper; the data were converted into electronic format manually.

Analyses

Traditional descriptive statistics of solution lengths relative to optimal are reported for the first two experiments. We used a novel method, reverse solution analysis, to investigate the solutions for bias toward ETF and DFC. ETF was operationalized here as the cumulative sum of all segments in the solution weighted by serial order of visitation. DFC is defined as the length of the first segment.

Reverse Solution Analysis

TSP solutions begin and end on the same city; they are closed loops. Therefore, a solution and its reverse form are equal in solution length (the only consideration by which solutions are evaluated in TSP). However, the two solutions may differ in terms of ETF or DFC (or other criteria of solution quality). We report bias toward a given consideration in the observed distribution when the percentage of solutions superior to their reverse forms with respect to that consideration exceeds that of the expected distribution in which both forms appear with equal frequency. The observed percentage over the expected percentage (50%) indicates the magnitude of the bias.

Because the PEBL TSP script automatically completes the last segment of solutions, these biases will be calculated for both the closed and open solutions, which omit the final segment returning to home. Note that Dataset 3 was generated using a paper and pencil format and required a return to home, and was included to show the extent to which automatic solution completion impacts performance.

Results and Discussion

Solution Length

Solution length provides a strong measure of overall efficiency (Figure 1). In Experiment 1, across all problem sizes, participants' mean solution lengths were 5.30% longer than optimal ($S.D.$ = 8.73%). Between set sizes, efficiency degraded with increasing problem size. Solutions to the 20-city problems showed the highest variance in solution length, an effect which was mirrored in the Experiment 2 results indicating that this likely reflects something about the city configurations for those problems.

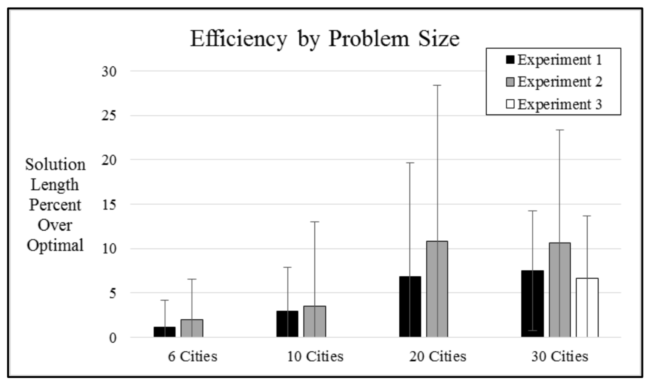

Figure 1: Efficiency by Problem Size for All Experiments. Error bars indicate standard deviation.

Experiment 2 efficiencies were largely consistent with those seen in Experiment 1. On average, participants' mean solution lengths were 6.76% longer than optimal ($S.D.$ = 12.67%). Between problem sizes, solution lengths increased with problem size, though efficiency in the 20- and 30-city problems was not notably different. Finally, participants' efficiency in solving the 30-city (M = 6.67%, $S.D.$ = 7%) problems in Experiment 3 was consistent with Experiment 1 performance for problems of the same size.

Other Solution Considerations

Of all Experiment 1 solutions (n = 1,358), 317 were optimal whereas 93 were designated poor (operationalized as 15% longer than optimal). Table 1 shows ETF and DFC bias by solution quality, measured using reverse solution analysis. Each block of cells indicates the percentage of solutions that favor that particular criterion given solutions of equal length.

The ETF bias block presents solutions quantified as either closed (complete, as generated by participants) or open (without the return to home). These results indicate the presence of ETF bias in the complete solutions. The magnitude of this effect appears consistent with that observed by Tenbrink and Wiener toward prioritizing larger versus smaller clusters earlier in the solution (2009; 58%). However, this effect is smaller when considering the solutions without the return to home. No clear trend in ETF bias was observed with respect to solution quality. DFC bias was also detected in these solutions, and the effects are likely related.

Table 1: Experiment 1 RSA Results, ETF and DFC Biases

ETF Bias

	Solution Quality	Percent Favoring ETF	Binomial Test Significance
Closed Solutions	**Optimal Solutions**	64.98	$p < .001$*
	All Solutions	63.40	$p < .001$*
	Poor Solutions	70.97	$p < .001$*
Open Solutions	**Optimal Solutions**	56.80	$p = .018$*
	All Solutions	52.30	$p = .098$
	Poor Solutions	45.16	$p = .417$

DFC Bias

	Solution Quality	Percent Favoring DFC	Binomial Test Significance
Closed Solutions	**Optimal Solutions**	67.19	$p < .001$*
	All Solutions	68.56	$p < .001$*
	Poor Solutions	80.65	$p < .001$*

To visualize this effect, we plotted the proportional distance of the solution covered by each segment in serial order (Figure 2). ETF and DFC biases are evidenced by shorter moves earlier in the solutions, or on the first move, respectively. Figure 2 shows reasonably uniform segment lengths for all except the final segment, indicating that most (but not all) of the bias effect appears to be explainable by a failure to account for the return to home cost.

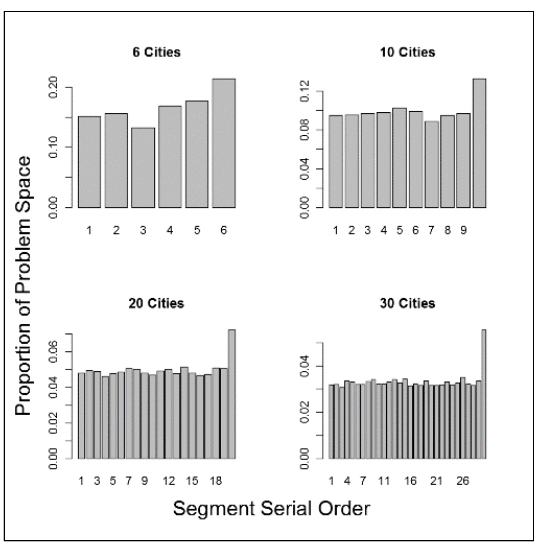

Figure 2: Proportion of problem space covered by each segment across Experiment 1 problems, by problem size.

The results of Experiment 1 show a robust bias toward solution forms that visit cities earlier in the solution at the expense of costs associated with the return to home.

In Experiment 2, participants only appeared to fail to account for the return to home on the larger 20- and 30-city problems, as indicated by their long final segment lengths (Figure 3). Aggregated across problem sizes, the results were similar to those seen in Experiment 1 with the exception that ETF bias disappeared entirely for the open solutions, and was not related to solution quality (Table 2).

Table 2: Experiment 2 RSA Results, ETF and DFC Biases

ETF Bias

	Solution Quality	Percent Favoring ETF	Binomial Test Significance
Closed Solutions	**Optimal Solutions**	59.84	$p < .001$*
	All Solutions	60.21	$p < .001$*
	Poor Solutions	60.74	$p = .008$*
Open Solutions	**Optimal Solutions**	51.71	$p = .539$
	All Solutions	51.21	$p = .378$
	Poor Solutions	46.01	$p = .347$

DFC Bias

	Solution Quality	Percent Favoring DFC	Binomial Test Significance
Closed Solutions	**Optimal Solutions**	55.64	$p = .031$*
	All Solutions	65.14	$p < .001$*
	Poor Solutions	75.46	$p < .001$*

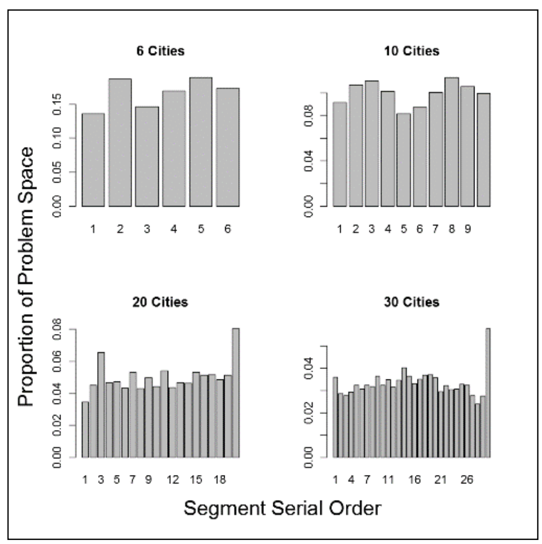

Figure 3: Proportion of problem space covered by each segment across Experiment 2 problems, by problem size.

Visualizing these data at a coarser grain size reveals ETF bias in the 6-city problems, and a trend toward it in the 10-city problems, with longer moves generally appearing in the second halves of the solution (Figure 4), despite no clear failure to return home.

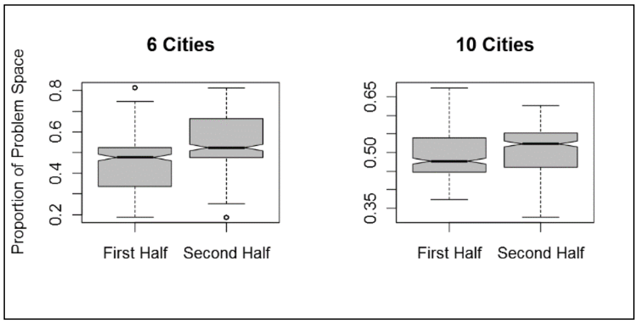

Figure 4: Distance covered by each half of the solution, Experiment 2. Lower values in first half indicate ETF bias.

One potential explanation for this effect is that it is an artifact of the experimental software. The PEBL TSP automatically completes solutions, so it is possible that participants' failure to account for the return home arises from the fact that they are not required to complete this section of the solution. Therefore, analysis of Dataset 3 tested for this effect in a paper and pencil version of TSP.

Dataset 3 consisted of 975 solutions to 9 30-city problems. 56.82% of these solutions favored the ETF-superior form (binomial test, $p < .001$) with 63.18% of solutions favoring the DFC-superior form (binomial test, $p < .001$). Figure 5 shows a strong failure to account for the return to home cost in all but one problem (Problem 3044).

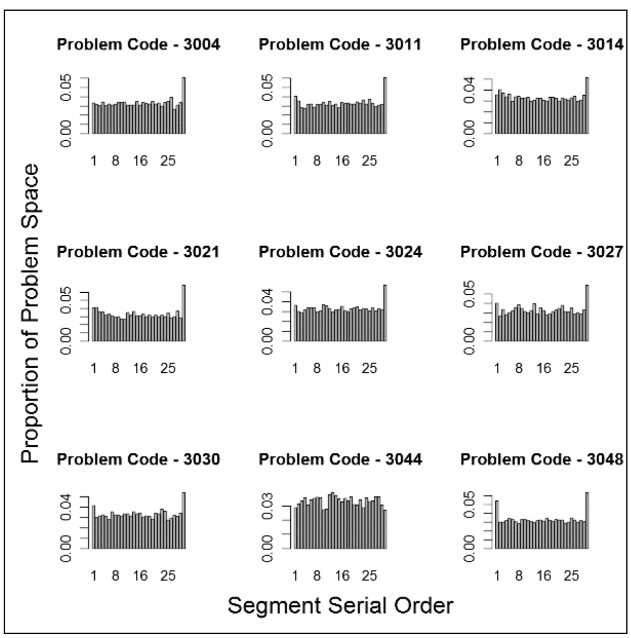

Figure 5. Proportion of problem space covered by each segment across each of the Dataset 3 problems.

To estimate the magnitude of the effect of the DFC bias and the failure to return home in each of these experiments, we divided the first and final segment lengths, respectively, by the average segment lengths (Table 3). One-way ANOVAs found significant effects of final / average segment length by solution quality for both Experiment 1, $F(1, 1356) = 94.85$, $p < .001$, and Experiment 2, $F(1, 953) = 165.1$, $p < .001$.

Table 3: First, Final / Average Relative Segment Length by Solution Quality, Mean Percent (*S.D.*)

	Solution Quality	**Experiment 1**	**Experiment 2**	**Dataset 3**
First Segment	**Optimal Solutions**	94.06 % (56.04 %)	85.86 % (52.82 %)	
	All Solutions	98.01 % (63.55 %)	88.45 % (62.53 %)	120.13 % (88.30 %)
	Poor Solutions	100.24 % (91.07 %)	101.89 % (96.91 %)	
Final Segment	**Optimal Solutions**	126.89 % (61.98 %)	98.03 % (60.89 %)	
	All Solutions	155.54 % (98.90 %)	148.41 % (111.09 %)	176.25 % (132.83 %)
	Poor Solutions	259.48 % (175.10 %)	226.37 % (171.83 %)	

Note: Solutions not aggregated by quality for Dataset 3 as optimal solutions to these problems were not available.

For Experiments 1 and 2, the average final segment length ranged from slightly shorter to over 2.5 times as long as the average segment length, with the paper and pencil TSP (Dataset 3) producing results falling somewhere in the middle. For Experiments 1 and 2, the final segment length generally increased as solution quality degraded, with the optimal solutions having much shorter final segment lengths relative to average than the poor solutions.

Similar effects were not observed for first / average segment lengths in Experiment 1, but the effect of solution quality on first / average segment lengths was observed in Experiment 2, $F(1, 953) = 6.89$, $p = .008$, with the better solutions producing shorter first segment lengths relative to average. First segment lengths in Dataset 3 were longer than average, though a causal mechanism is not readily apparent.

Finally, in Experiments 1 and 2 the larger problem sizes generally produced longer first and final segment lengths relative to average (see Table 4). One way ANOVAs revealed significant effects of problem size on first, $F(1, 1048) = 11.79$, $p < .001$, and final, $F(1, 1048) = 102.60$, $p < .001$, segment lengths in Experiment 2. The effect of problem size on segment length was observed for the final, $F(1, 1356) = 20.66$, $p < .001$, but not first, $F(1, 1356) = 0.07$, $p = .799$, segment lengths in Experiment 1.

Table 4: First, Final / Average Relative Segment Length by Problem Size, Mean Percent (*S.D.*)

	Problem Size	**Experiment 1**	**Experiment 2**
First Segment	**6**	96.97 % (71.42 %)	83.61 % (57.00 %)
	10	97.49 % (59.66 %)	93.44 % (49.89 %)
	20	98.26 % (68.03 %)	69.56 % (61.06 %)
	30	96.06 % (62.04 %)	110.68 % (81.85 %)
Final Segment	**6**	145.34 % (77.70 %)	113.43 % (77.11 %)
	10	143.07 % (78.29 %)	102.76 % (58.71 %)
	20	152.12 % (102.05 %)	175.07 % (141.70 %)
	30	175.39 % (129.40 %)	183.84 % (137.20 %)

Human Results Summary and Discussion

The results presented above demonstrate, for the first time to our knowledge, the presence of considerations pertinent to naturalistic TSP-like tasks in traditional TSP solutions. Participants (1) produced solutions that reduce distance to first contact and (2) preferred visiting locations early in the solution at the expense of overall solution length, resulting in higher distance-discounted reward. However, (3) this effect is driven largely by a preference for solution forms with a longer return to home. (4) This effect is robust to the test delivery format (i.e., computer with an automatic return to home versus manual pen and paper format) and (5) the magnitude of this bias is related to the quality of the closed solutions – better solutions reduce the discrepancy between final and average segment lengths.

Points 1 and 2 suggest that humans solve a more general problem than TSP task instructions would require; in addition to solution length, humans consider ETF and DFC during problem solving. And, in light of the constraint of point 3, they seem to do so at the expense of task performance, though better human solutions tend to consider the problem space more globally, therefore accounting for the return home.

Subject matter expert interviews (Perelman, 2015) indicate that ETF and DFC are critical for certain tasks, and prior research (Perelman & Mueller, 2015) has shown that humans can adapt their solution criteria to fit specific tasks. However, the results of the present study show that even when tasked with minimizing solution length in a traditional TSP, humans still generate solutions that account for considerations relevant in naturalistic spatial problem solving tasks. This suggests a common mechanism used for both TSP, and for naturalistic TSP-like tasks. A simple computational model was developed by Perelman (2015) to describe adaptive behavior in TSP-like problems, which we apply here to investigate an underlying adaptive mechanism capable of producing the above effects in TSP.

Modeling

In light of the human results, we used a computational model designed to permit this flexibility in strategic control that is capable of solving TSPs using limited at-a-glance information about the problem space. The goal of this model was to use a scheme capable of adapting to task requirements (i.e., it incorporates strategic considerations such as ETF into solution planning) to reproduce human efficiency dynamics and solution characteristics, specifically the bias toward ETF-superior solution forms.

The model uses a two-layer hierarchical structure: a computationally inexpensive local decision making algorithm (nearest neighbor) guided by a general *plan* (see Figure 6) that considers multiple criteria that can be tailored to specific goals and tasks (i.e., path length minimization, as in TSP, versus discounted-rewards used in naturalistic tasks). This higher level plan representation consists of a small number of segments; the model solves for all the cities within each segment in sequence. The plan is initially drawn by running K-means clustering over the problem space (k = 6) then connecting the cluster centroids from the starting position by nearest neighbor.

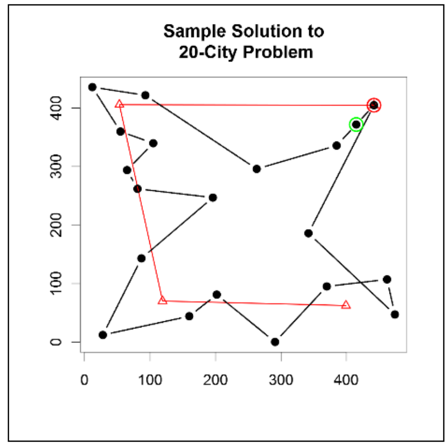

Figure 6. The higher level plan (red segments) guides local solutions (black lines) among the cities (black dots) from the start location (red; first move is green to show direction)

The plan is then iteratively fit to the data by minimizing a cost function comprised of a linear combination of five weighted parameters, (1) log number of segments, intended to represent plan complexity, (2) goodness of fit, the average distance between a plan segment and its constituent cities, (3) plan length, (4) distance-discounted reward, the sum of the path lengths of all segments discounted by their serial order, and (5) the average angle between segments, intended to penalize doubling back. This plan is fit to the data using 500 iterations during which a point in the plan is added, deleted, moved, or swapped in serial order with another. This optimization process was not intended to duplicate that used by humans to solve the problem, but rather to demonstrate that a model that incorporates multiple criteria can account for some patterns in the human data.

Modeling Results

The present model solved the 10-, 20-, and 30-city TSP problems used in Experiment 1 20 times each, and the solutions were analyzed using the methods described above.

Solution Length

To evaluate model solution efficiency, solution lengths were compared to those of the optimal solutions. Across all problem sizes, model solutions lengths were 16.08% longer than optimal (*S.D.* = 11.56%). Interestingly, as with the human subjects, the model produced the greatest variance in solving the 20-city problems (Figure 7), indicating that human performance on these problems can be attributed to properties of the problem spaces rather than a fluke in our particular sample. While the model was less efficient than humans, it displayed similar dynamics in the present problem set for the change in solution quality at the experimental problem sizes, and variance within set sizes.

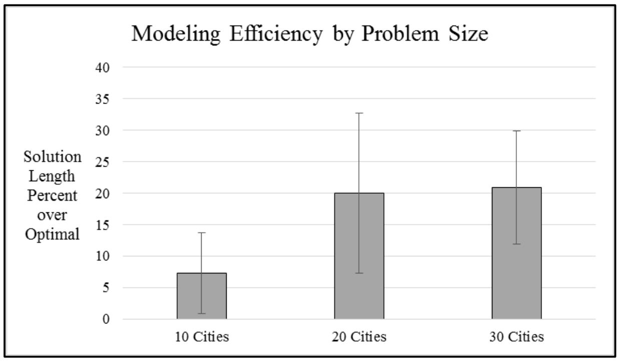

Figure 7. Model solution length by problem size. Error bars indicate standard deviation.

The second goal for our model was to replicate the human bias toward DFC- and ETF-superior solution forms. We investigated the model solutions using the method applied to the previous experiments and found that the model favored ETF- and DFC-superior solutions more strongly than humans (Table 5), preferring a biased solution in nearly every case except for the 20-city problems. Figure 8 demonstrates qualitatively the effects of ETF and DFC bias in the model's solutions, along with a failure to account for the return to home producing segments roughly twice as long as average.

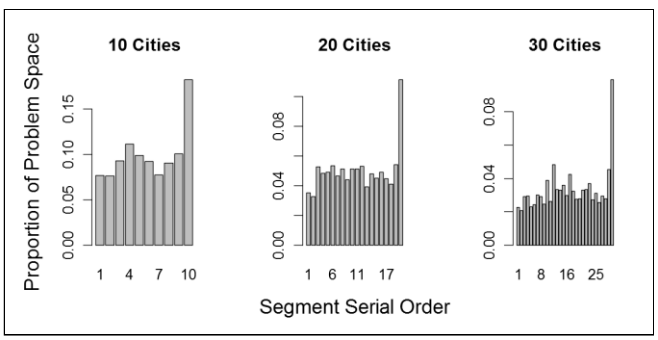

Figure 8: Proportion of problem space covered by each segment across Experiment 1 problems, by problem size.

As with the human solutions, we further quantified the ETF and DFC biases, and the effect of the failure to return home, by comparing the first and final segment lengths, respectively, to the lengths of the average segments on those trials. The model produced solutions with values (Table 5) similar to those of the poor human solutions (Table 3), including comparatively short first segment lengths, and long final segment lengths (i.e., the return to home) that increased with problem size. However, unlike the human solutions, the model produced solutions with first segment lengths that decreased with increasing problem size relative to the average segment lengths on those trials. Finally, the model produced solutions with shorter relative first segment lengths, but longer final relative segment lengths, compared to the human solutions.

Table 5: Final / Average Segment, Mean Percent (*S.D.*), and Proportion of Solutions Favoring Each Bias

	10 Cities	**20 Cities**	**30 Cities**
First / Average	79.16 %	70.85 %	67.95 %
Segment Length	(41.91 %)	(31.93 %)	(25.90%)
Final / Average	205.85 %	257.52 %	327.16 %
Segment Length	(56.19 %)	(114.72 %)	(138.09 %)
Percent Favoring ETF	99.33 %	89.67 %	100 %
Percent Favoring DFC	100 %	94.33 %	100 %

Last, we investigated a potential criticism of the present model – that a nearest neighbor model would be equally efficient. We compared model and human solution lengths to those generated using nearest neighbor. The present model produced solutions to the 10-city problems that were 4.1% shorter than nearest neighbor, equal in length for the 20-city problems, and 1.1% shorter for the 30-city problems. Human solutions in Experiment 1 generally showed a similar pattern (6%, 10% and 10%, respectively), though the poor human solutions were on average 3-9% longer than those produced by nearest neighbor. In summary, the model solutions were less efficient than human solutions on average, but more efficient than the nearest neighbor and poor human solutions.

General Discussion

The behavioral results of the present study show, for the first time, that human TSP solutions consider ETF and DFC, criteria that are irrelevant to TSP, but critical in the real world. This manifested here as a failure to account for the return to home, and the magnitude of this bias was related to solution quality – poor solutions had longer final segments.

A model that adjusts a linear plan to fit the problem space according to a number of criteria related to TSP, real-world TSP-like problems, and plan complexity, exhibits behavior similar to humans in this task. We expect that efficiency could be greatly improved via dynamic re-planning to match the human data. In adapting the present model, the agent would adjust its higher level plan after solving for the points within each segment in serial order. In this way, a parameter estimated from prior eyetracking studies (e.g., Kong & Schunn, 2007) would govern the iterations spent in dynamic replanning.

Taken together, the results of the present study hold implications for modeling human performance in spatial combinatorial optimization problems. Specifically, the results speak to the importance of granularity and sequence in representing the problem space. Many algorithms, such as those implementing a convex hull, solve the problem exhaustively at presentation. The behavioral and modeling results presented here are consistent with prior work suggesting that humans approach these problems using a mechanism that provides a means of solving the problem efficiently without the mental burden of generating and maintaining an exhaustive solution in memory, at the expense of efficiency later in the route. Finally, this mechanism is consistent with producing solutions to discounted-reward problems that are more common than path length optimization in naturalistic tasks.

References

Antoniades, A., Kim, H. J., & Sastry, S. (2003, December). Pursuit-evasion strategies for teams of multiple agents with incomplete information. In *Proceedings of the 42nd IEEE Conference on Decision and Control, 1,* 756-761.

Blum, A., Chawla, S., Karger, D. R., Lane, T., Meyerson, A., & Minkoff, M. (2007). Approximation algorithms for orienteering and discounted-reward TSP. *SIAM Journal on Computing*, *37*, 653-670.

Chronicle, E. P., MacGregor, J. N., Lee, M., Ormerod, T. C., & Hughes, P. (2008). Individual differences in Optimization Problem Solving: Reconciling Conflicting Results. *The Journal of Problem Solving*, *2*(1), 41-49.

De Jong, L. W., Gereke, B., Martin, G. M., & Fellous, J. (2011). The traveling salesrat: Insights into the dynamics of efficient spatial navigation in the rodent. *Journal of Neural Engineering, 8.* doi:10.1088/1741-2560/8/6/065010.

Dry, M., Lee, M. D., Vickers, D., & Hughes, P. (2006). Human performance on visually presented traveling salesperson problems with varying numbers of nodes. *Journal of Problem Solving, 1,* 20-32.

Dantzig, G. B. & Ramser, J. H. (1959). The truck dispatching problem. *Management Science, 6,* 80-91.

Kong, X. & Schunn, C. D. (2007). Information seeking in complex problem solving. In *Proceedings of the 8th International Conference on Cognitive Modeling*, Oxford: UK, 261-266.

Krolak, P., Felts, W., & Marble, G. (1971). A man-machine approach toward solving the Traveling Salesman Problem. *Communications of the ACM, 14,* 327-334.

MacGregor, J. N. & Ormerod, T. C. (1996). Human performance on the Traveling Salesman Problem. *Perception & Psychophysics, 58,* 527-539.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. (2000). A model of human performance on the traveling salesperson problem. *Memory & Cognition, 7,* 1183-1190.

Mueller, S. T. (2014). PEBL: The Psychology experiment building language (Version 0.14) [Computer experiment programming language]. Retrieved June 2014 from http://pebl.sourceforge.net.

Mueller, S. T., Perelman, B. S., Tan, Y., & Thanasuan, K. (2015). Development of the PEBL Traveling Salesman Problem Computerized Testbed. *The Journal of Problem Solving*, *8*, 4.

Mueller, S. T. & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and the PEBL Test Battery. *Journal of Neuroscience Methods, 222,* 250-259. Doi: 10.1016/j.jneumeth.2014.10.024.Perelman, B. S. & Mueller, S. T. (2013). A Neurocomputational approach to modeling human performance in simulated unmanned aerial search. In *Proceedings of the 12th International Conference on Cognitive Modeling*, Ottawa, CA.

Perelman, B. S. (2015). A naturalistic computational model of human behavior in navigation and search tasks. *Doctoral Dissertation*.

Perelman, B. S. & Mueller, S. T. (2015). Identifying mental models of search in a simulated flight task using a pathmapping approach. In *Proceedings of the 18th International Symposium on Aviation Psychology*, OH.

Pizlo, Z., Stefanov, E., Saalweachter, J., Li, Z., Haxhimusa, Y., & Kropatsch, W. G. (2006). Traveling salesman problem: A foveating pyramid model. *The Journal of Problem Solving*, *1*, 8.

Ragni, M. & Wiener, J. M. (2012). Constraints, Inferences, and the Shortest Path: Which paths do we prefer? In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Sapporo, Japan.

Tenbrink, T. & Seifert, I. (2011). Conceptual layers and strategies in tour planning. *Cognitive Processes, 12,* 109-125.

Tenbrink, T. & Wiener, J. (2009). The verbalization of multiple strategies in a variant of the traveling salesman problem. *Cognitive Processes, 10,* 143-161.

Wiener, J. M., Schnee, A., & Mallot, H. A. (2004). Use and interaction of navigation strategies in regionalized environments. *Journal of Environmental Psychology*, *24*, 475-493.