On the Separability of Social Navigational Behaviors in Virtual Reality

Samuel S. Sohn[1] (samuel.sohn@rutgers.edu), Serena DeStefani[3], Mathew Schwartz[4], Jacob Feldman[1], Mubbasir Kapadia[2], Karin Stromswold[1]
[1]Department of Psychology & Center for Cognitive Science, Rutgers University – New Brunswick
[2]Department of Computer Science, Rutgers University – New Brunswick
[3]Department of Psychiatry & Behavioral Health, The Ohio State University
[4]New Jersey Institute of Technology

Abstract

Human navigation is shaped by cognitive strategies, spatial awareness, and learned heuristics, yet existing models struggle to capture individual differences in wayfinding. To investigate the cognitive basis of navigational behavior, we conducted a virtual reality experiment where participants maneuvered around a human obstacle in a controlled, static environment. Using trajectory-based features, we classified participants with PartNet, a neural network that outperformed ElasticNet and Random Forest classifiers. While PartNet captured subtle yet consistent behavioral patterns, its interpretability was limited. To address this, we developed an analysis pipeline revealing key behavioral factors, showing that navigational styles differ primarily in midline adherence and speed. Clustering and embedding analyses further demonstrated participant separability, highlighting both individual distinctions and shared tendencies. By identifying structured variability in navigation, our work advances cognitive models of spatial decision-making, informing theories of wayfinding, predictive modeling of human movement, and applications in assistive navigation and urban design.

Keywords: navigation, modeling, data visualization, interpretability

Introduction

Modeling human navigation is crucial for urban planning and the safe management of public events (Koumetio Tekouabou, Diop, Azmi, Jaligot, & Chenal, 2022; Casali, Aydin, & Comes, 2022; Sohn et al., 2020). A key challenge is accounting for navigation around social obstacles, which, unlike physical barriers, involve complex social rules and individual differences (Kim et al., 2015). Traditional particle-based models fail to capture these nuances (Vizzari, Crociani, & Bandini, 2020), such as the varying acceptable distances when passing in front of or behind a person. Addressing this gap requires a deeper understanding of how people navigate in dynamic environments, balancing shared movement tendencies with individual differences.

Identifying whether navigational behaviors cluster into common styles or remain highly individualized has significant implications. If behaviors generalize across individuals, predictive models can be simplified and applied broadly. Conversely, strong individual differences may necessitate more personalized approaches. To explore this, we conducted a controlled virtual reality experiment to analyze fine-grained trajectory features and classify participants based on their navigation tendencies.

We introduce PartNet, a neural network that surpasses ElasticNet and Random Forest classifiers in distinguishing participants while capturing subtle yet consistent navigational tendencies. However, the strength of complex models often comes at the cost of interpretability, limiting their practical application. To bridge this gap, we designed an analysis pipeline that identifies key behavioral features, revealing how individuals differ in midline adherence, speed, and obstacle avoidance strategies. By focusing on interpretability, we aim to connect machine learning-driven insights with real-world applications in public health policy and urban planning. This allows us to articulate the factors influencing navigation, justify behavioral clusters transparently, and translate findings into practical guidelines (Burkart & Huber, 2021; Carvalho, Pereira, & Cardoso, 2019).

Related Work

Proxemics theory, introduced by Edward T. Hall, examines how humans use and perceive space (Hall, 1968). Hall identified concentric zones around individuals, each serving different social functions: public, personal, and intimate areas (Watson, 2014). Subsequent research has revealed cultural variations and contextual factors influencing these spatial preferences (Cristani et al., 2011). The model was later refined to account for asymmetry, considering differences between the front and back of the body (Hayduk, 1981), as well as the dominant and non-dominant hands (Gérin-Lajoie, Richards, Fung, & McFadyen, 2008). Recent advancements have expanded on Hall's model, introducing concepts such as activity spaces and affordance areas (Schaumann, Sohn, Usman, Haworth, & Kapadia, 2019; Frohnwieser, Hopf, & Oberzaucher, 2013), providing a more nuanced understanding of spatial dynamics in social interactions.

Proxemics plays a crucial role in social navigation, which studies how individuals move through environments populated by others, who can be in turns sources of information to approach (Dalton, Hölscher, & Montello, 2019) or obstacles to avoid (Crociani, Vizzari, & Bandini, 2018; Haghani & Sarvi, 2017) . This field has significant implications for urban planning (Farr, Kleinschmidt, Yarlagadda, & Mengersen, 2012), socially aware robotics (Rios-Martinez, Spalanzani, & Laugier, 2015), and crowd simulation (Daza, Barrios-Aranibar, Diaz-Amado, Cardinale, & Vilasboas, 2021). Traditional proxemics studies are limited by their focus on static, fixed areas. In contrast, model-based approaches offer a more dynamic representation of social navigation (Helbing

& Molnár, 1995; Corbetta & Toschi, 2023; Yue, Manocha, & Wang, 2022) but often overlook individual differences (Zheng, Zhong, & Liu, 2009).

Data-driven clustering techniques present a promising alternative (Alahi et al., 2017), capturing real-world behavioral patterns while accounting for individual variability (Antonenko, Toy, & Niederhauser, 2012). Clustering methods offer several advantages: they reflect authentic human behavior (Du, 2010) rather than theoretical assumptions and they are flexible, adapting to specific contexts and environments. In addition, they can identify common patterns across diverse individuals, and they can potentially reveal unexpected navigation strategies not accounted for in existing models (Atev, Miller, & Papanikolopoulos, 2010). Supervised methods in particular can help capture more nuanced patterns in trajectory data (Bian, Tian, Tang, & Tao, 2018). By leveraging actual trajectory data, clustering techniques provide a more nuanced and realistic foundation for understanding and predicting social navigation behaviors.

Field experiments in social navigation, while offering strong ecological validity, often present challenges in controlling environmental variables and comparing conditions due to inherent noise. These limitations are an obstacle to the extraction of clear, generalizable insights. Laboratory-based experiments offer an alternative, providing greater control over variables (Haghani, 2020). In addition to greater control over environmental conditions, VR experiments offer several key advantages, like the ability to rapidly present diverse scenarios, and to consistently replicate experimental conditions across participants (Haghani & Sarvi, 2018). The recent proliferation of commercial virtual reality (VR) headsets has significantly increased the feasibility and efficiency of conducting virtual social navigation experiments. Research has demonstrated that human behavior in VR environments closely mimics real-world behavior, supporting the validity of VR-based studies (Haq, Hill, & Pramanik, 2005; Kinateder & Warren, 2016; Dong et al., 2022).

In this study, we examined the impact of environmental factors, such as mask usage and perceived safety, on trajectory quality. Given that a primary challenge of machine learning clustering techniques is their lack of interpretability (Doshi-Velez & Kim, 2017), we prioritized this aspect throughout our evaluation process.

Dataset

We conducted a comprehensive human behavioral experiment in virtual reality to investigate human navigational behaviors under various conditions. The study involved 33 participants (17 females and 16 males), all of whom were undergraduate students at a university whose Institutional Review Board (IRB) approved the study protocol. Four individuals were excluded from the analysis: three for not adhering to instructions and one due to incomplete data.

The participants were tasked with navigating through a virtual room (wide: -2.25m $\leq$ x $\leq$ 2.25m, length: 0m $\leq$ y $\leq$ 6m) by maneuvering around an obstacle at (0, 3) to reach a door on the opposite end at (0, 6) (Figure 1). To ensure a thorough exploration of factors influencing navigation behavior, we implemented a full factorial design with four trial conditions:

- **Obstacle type**: human standing and wearing surgical mask, human standing unmasked, human sitting in a chair masked, human sitting in a chair unmasked.
- **Obstacle orientation**: the 4 cardinal directions, where north points toward the exit.
- **Environment safety**: following or not following COVID-safety protocols.
- **Participant safety**: wearing or not wearing a face mask.

Each participant completed 64 trials composed of all combinations of these conditions (4 obstacle types $\times$ 4 orientations $\times$ 2 environment safety conditions $\times$ 2 participant safety conditions) in approximately 20 minutes. This experimental setup allowed us to collect rich trajectory data for the analysis of human movement patterns in constrained environments.

During each trial, we recorded the position of the participant's VR headset at 50Hz as they moved through the environment and smoothed out lateral head movement to reflect the movement of their center of mass (Kavanagh, Morrison, & Barrett, 2005). To further improve data quality, we truncated the start of each participant's data below the threshold y = 1m to remove noise associated with participants getting situated at the beginning of each trial (Figure 1). We also removed trials exceeding a duration threshold of 10 seconds, reducing the total number of trials from 1,856 (64 trials $\times$ 29 participants) to 1,834. This criterion was based on the average room-crossing time of 6-7 seconds. Longer trials typically resulted from participants deviating from instructions to walk naturally, instead becoming distracted or interacting with VR artifacts.

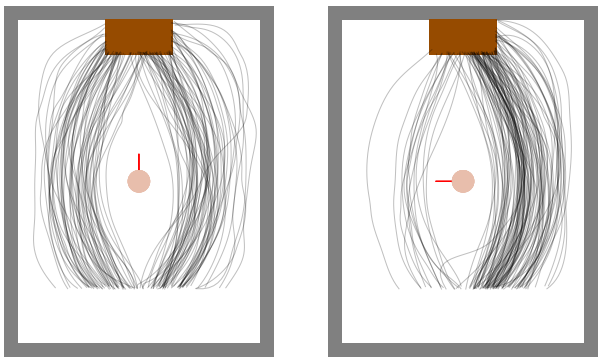

Figure 1: The above figure shows experimental setup in VR. The participant needs to move around the obstacle (beige) facing the red line to reach the brown door.

Feature Extraction

The 64 trial conditions were encoded as 6 *condition features*: 2 binary values for whether the human obstacle was sitting

and whether it was masked (1,2), 2 values representing the orientation of the obstacle as a unit vector (3,4), and 2 binary values for environment safety and participant safety (5,6). From the position data, we extracted a comprehensive set of *navigation metrics* to characterize each participant's behavior: (7) average distance of the participant from the midline x = 0 (ADFM), (8) maximum distance from the midline (MDFM), (9) horizontal distance from the obstacle at y = 3, (10) average speed, (11) maximum speed, (12) trajectory duration, (13) whether the participant moves left or right to avoid the obstacle, and (14) average angular velocity. Since many of the trajectories were observed to follow a Gaussian curve along the midline, the parameters of a Gaussian distribution were fitted to each trajectory, from which the mean (15) and the standard deviation (16) were used as additional metrics. Finally, we added a metric for the closest the participant gets to the obstacle (17) as well as a direction alignment metric (18), which outputs 1 when the participant chooses to move in front of the obstacle facing east or west, -1 when the participant moves behind the obstacle facing east or west, and 0 when the obstacle is facing north or south. This rich set of features captures both spatial and temporal aspects of the participants' movement, with a focus on understanding proxemics behavior and avoidance patterns.

Participant Classification

Our analysis on the separability of human navigational behaviors begins with modeling the participant classification problem, where we predict which participant produced a set of navigational metrics given the trial conditions. More formally, the input consists of 18 features (i.e., the 12 navigation metrics and 6 condition features), and the output is represented as a 29-dimensional vector, where the correct participant's dimension is set to 1 and all others are 0.

We consider three models for learning to classify participants: ElasticNet (Zou & Hastie, 2005), the Random Forest Classifier (RFC) (Breiman, 2001), and PartNet —our proposed neural network. ElasticNet was chosen to represent linear regression models (over logistic regression), because it penalizes the learned coefficients in 2 ways (Lasso (Tibshirani, 1996) and Ridge (Hoerl & Kennard, 1970)) to improve the predictive performance and interpretability of linear regression models. By combining these penalties, ElasticNet addresses the individual limitations that Lasso and Ridge regression have in dealing with multicollinearity (e.g., between ADFM and MDFM or average and maximum speed) and performing feature selection respectively. The RFC complements Elastic Net by addressing its limitations in modeling nonlinear relationships. It leverages an ensemble of decision trees to capture intricate interactions and dependencies among features, while still offering interpretability through feature importance scores.

In contrast to these more interpretable models, we propose PartNet, a simple neural network comprising of linear layers that learn relationships between features and can change their dimensionality, BatchNorm layers that stabilize training (Ioffe & Szegedy, 2015), ReLU layers that introduce nonlinearity (Agarap, 2018), and a Softmax layer to ensure that the output is a probability distribution over the 29 participants. Figure 2 illustrates how these layers are organized in the PartNet architecture and includes their output dimensions. The blue highlighted embedding layer is an intermediary feature representation within PartNet, which we later extract for analysis. PartNet is trained for 50 epochs using stochastic gradient descent to optimize its parameters with a learning rate of 0.01, momentum value of 0.9 (Sutskever, Martens, Dahl, & Hinton, 2013), and a weight decay of 0.001 to prevent overfitting by penalizing large weights. The model used Cross-Entropy Loss to quantify the difference between the predicted participant probabilities and the ground truth participant labels.

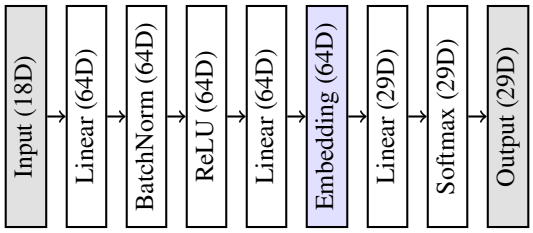

Figure 2: The above diagram is of PartNet's architecture.

In order to train ElasticNet, the RFC, and PartNet, the 1,834 data points were randomized 30 times and for each randomization, an 80-20 training-testing split was made. While the condition features were already normalized, the navigation metrics had to be standardized according to the training set of each randomized split. Figure 3 shows the traditional accuracy of each model trained on all features (All Feat.) at $k = 1$, where only a model's top guess is evaluated. Looking at this result alone, it appears that RFCs outperform ElasticNet and PartNet on average, but this comes with an exorbitant standard deviation. On the other hand, PartNet is significantly better than the RFC in consistency and ElasticNet in accuracy.

Looking beyond the traditional top guess of each model, we observe an interesting linear signature to the RFC's accuracy curve as the number of guesses k increases. It appears that the RFC's learned features are shallow and overfitted, because its accuracy increases linearly with k much like random chance. In contrast, PartNet and ElasticNet exhibit asymptotically greater increase in accuracy, which permanently overtake the RFC after $k = 2$ guesses for PartNet and $k = 4$ guesses for ElasticNet. Put another way, to achieve $\sim$70% accuracy, PartNet requires 3 guesses, ElasticNet requires 5 guesses, and the RFC requires 9 guesses on average. PartNet and ElasticNet share similar logarithmic accuracy signatures, but PartNet is consistently more performant. This signature (and how it differs from the linear signatures) is informative of the separability of participants' navigational behaviors at a coarse level. The large initial leap in accuracy suggests that there are navigational behaviors that are shared among some par-

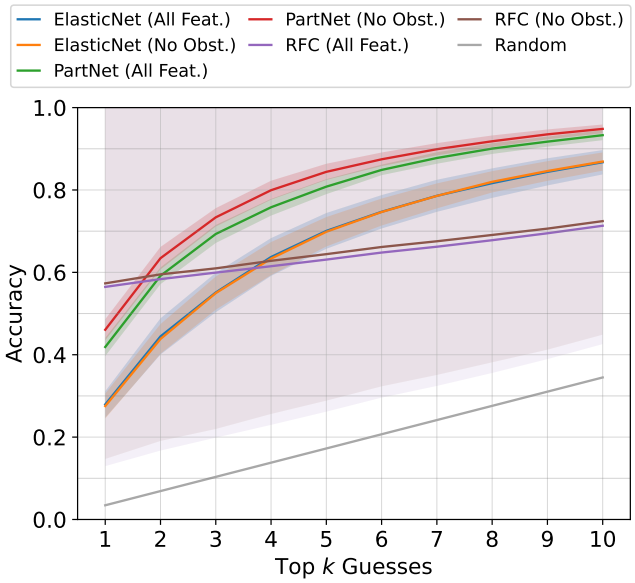

Figure 3: The above plot shows the accuracy of ElasticNet, the RFC, and PartNet as a function of the top-k guesses with and without obstacle-related features.

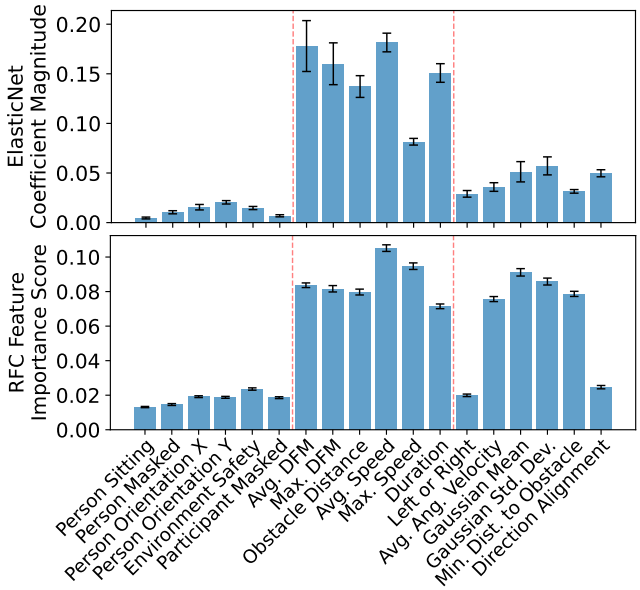

Figure 4: The above bar plots shows the magnitude of ElasticNet's coefficients and the RFC's feature importance scores. Features with higher values are more important for classifying participants.

ticipants. *The fewer the people that share a behavior, the earlier (in terms of guesses) it will yield an improvement in accuracy.* If all behaviors were unique (i.e., maximally informative), a model with enough predictive power would achieve 100% for all k, and if all behaviors were shared (or totally random, which are both minimally informative), the model accuracy would look like random chance. The two signatures (linear and logarithmic) lie on the continuum between these theoretical extremes. We assert that the RFC does not have enough predictive power to learn meaningful *similarities* between participants (hence the linear growth), though it learns some meaningful *differences* on average (hence the high initial accuracy).

Feature Analysis

Figure 4 shows that ElasticNet and the RFC similarly undervalue condition features (Figure 4, left) as well as directional features (left/right and alignment), and they both value distance and speed-based metrics (Figure 4, middle). However, the RFC finds more value in metrics that inform on trajectory shape (Figure 4, right) with more fidelity such as angular velocity (which encodes some curvature) or the fitted Gaussians' parameters. This aligns with our prior assertion that it tries to learn differences better.

Since there are no readily interpretable features within PartNet, we perform an ablation study, where we train PartNet with 11 subsets of the original 18 features. For conciseness, we refer to *omitted* features using the numbers (1-18) from the **Feature Extraction** Section. The features subsets are as follows: no person sitting (1), no person masked (2), no person orientation (3-4,18), no environment safety (5), no participant safety (6), no trial conditions (1-6,18), no obstacle information (1-4,18), no distance from midline or DFM (7-9), no speed (10-12), no curvature (14,16-17), and no direction (13,18).

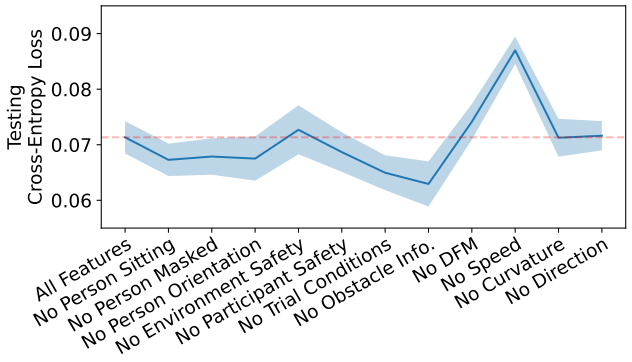

Figure 5: The above plot shows the result of the PartNet ablation study. Some features cause overfitting (e.g., obstacle information), while others play a crucial role to accuracy (e.g., distance and speed information).

If all features are productive toward learning, we expect each ablation to increase loss. Figure 5 shows some agreement with ElasticNet and the RFC in that distance- and speed-based features are crucial, as indicated by the increases in loss for No DFM and No Speed. However, it appears that some features are highly counterproductive, namely condition features carrying obstacle information (1-4), which significantly *improve* performance when removed from PartNet. Environ-

ment safety was the only condition feature that did not degrade PartNet's performance on average. This suggests that for PartNet, *how* participants move (as captured by navigation metrics) helps to classify participants more robustly than *why* they move (based on the obstacle) —a distinction that is shared by other models but not treated as extremely as PartNet. We posit that this is due to the uniformity in participants' treatment of the obstacle (Figure 1) and nonuniformity in their treatment of the environment's COVID-related safety (Moran et al., 2021).

Based on this finding, we removed obstacle information from each model and reevaluated their accuracy as a function of k (Figure 3). ElasticNet was largely unaffected by this ablation, because it already weighs condition features lowly. There was some visible effect on the RFC in that the mean accuracy increases marginally but consistently across all k. The most pronounced improvement was observed with PartNet, which remains the most performant model.

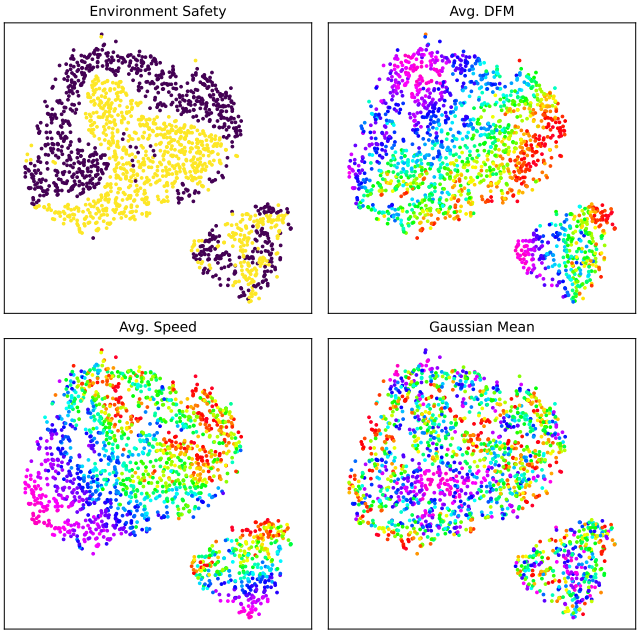

Figure 6: The above plots show different metrics mapped onto PartNet's embedding space after dimensionality reduction using t-SNE.

Embedding Analysis

The ablation study is a black-box evaluation method that is useful, but there is more to gain from analyzing the feature representations within PartNet. While learning to classify participants, PartNet creates a feature representation that tries to separate participants' behaviors into different regions. However, as aforementioned, a perfect separation is likely impossible due to shared behaviors. In PartNet's architecture, layers closer to the output have feature representations that are better separated. Therefore, we extract features (colored blue in Figure 2) before they enter the final linear layer, which is a participant embedding.

It is difficult to interpret the individual dimensions of the embedding, but we can determine which known features they are correlated with by first using t-SNE (Van der Maaten & Hinton, 2008), a dimensionality reduction technique, to project the embeddings onto a 2D plane (Figure 6). This creates a human-interpretable view of the aforementioned "participant regions" that has been parameterized (with perplexity = 120) to preserve broader structures among the embeddings. In theory, if a particular metric is important to PartNet, it will influence the broad structure the embeddings and there should be some observable continuity when plotting it onto this 2D view.

Figure 6 shows environment safety, ADFM, average speed, and the Gaussian means visualized over the 2D view. It is clearly evident that the first three features are preserved in the embeddings. Environment safety exhibits a clear boundary between safe and unsafe behaviors that creates complex interactions in the way that it cuts through ADFM and average speed. Interestingly, ADFM and average speed, which all 3 models consider to be the most important, are continuous along orthogonal axes in the 2D view. The Gaussian mean, which is considered less important, appears to increase from the centers of the 2 clusters to the perimeters (i.e., the axis orthogonal to both ADFM and average speed, coming out of the page). These results are in full alignment with the feature importance scores and ablation study (Figures 4 and 5).

Clustering Analysis

Features embed more about participants as they pass through PartNet. Accordingly, the early stage ablation study offers feature-centric insights, and the mid-stage embedding analysis bridges features and participants. We now conduct a late stage clustering analysis to fully connect this pipeline. For each test case, PartNet outputs a probability distribution over participants, which implicitly encodes which participants it is confused about and to what degree. All of these distributions are aggregated into the confusion matrix in Figure 8, which has been averaged cell-wise across all 30 training-testing splits. We apply Louvain community detection on the matrix (treating it like an adjacency graph) to cluster participants based on how much they were confused with each other (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008; Brusco & Steinley, 2006). Table 1, rows 3 and 6 list the 14 detected clusters. Looking at the confusion matrix's diagonal, it is evident that some participants' values are high (indicating low confusion), while others' are low (indicating high confusion). We quantify this confusion more accurately using the Shannon entropy of each row (Shannon, 1948) (Table 1, rows 2 and 5).

This entropy is precisely what measures the titular separability of participants. This is illustrated well by clusters 11 and 13. According to the high entropy of participants in cluster 11, we expect their regions in the embedding space to be spread out, overlapping other participants' regions. Since

Participant	5	10	20	21	7	11	16	3	24	12	29	15	26	2	6
Entropy	3.0	2.5	3.5	2.8	3.5	3.9	3.4	3.5	3.4	1.6	1.8	3.0	2.0	3.4	2.9
Cluster	10	10	10	10	11	11	11	3	3	13	13	12	12	2	6

Participant	1	8	13	9	18	28	4	27	14	23	17	22	19	25	
Entropy	3.9	3.7	4.1	3.6	3.5	3.9	3.9	3.7	4.1	4.1	3.6	2.7	3.4	2.5	
Cluster	8	8	8	9	9	9	4	4	14	14	5	5	1	7	

Table 1: This table shows which cluster each participant belongs to and the entropy of each participant's row in the averaged PartNet confusion matrix.

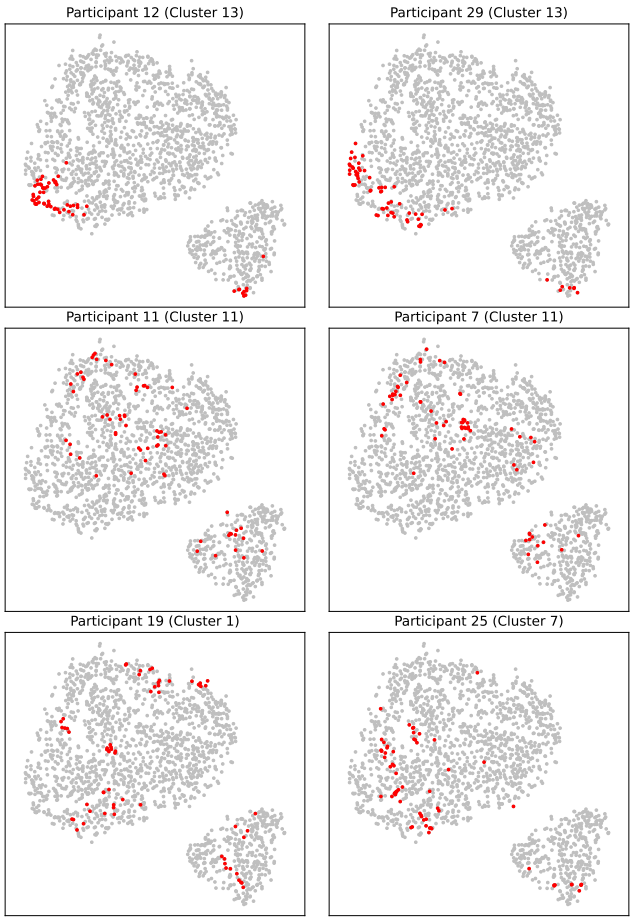

Figure 7: Plots (a-d) show the points in the embedding space corresponding to particilar participants. Participants belonging to the same cluster appear to have similar distributions.

they belong to the same cluster, we expect them to overlap with others in similar ways. On the contrary, we have cluster 13, whose participants have low entropy and thereby (in theory) compact regions that are well-separated from others, but close to each other. Figure 7 aligns completely with our expectations. Given the commonality of cluster 11's behaviors, we expect that their widespread regions match up with more average distance- and speed-based features, and that cluster 13's regions have much more extreme features values. Figure 6 strongly supports these expectations. Some clusters only

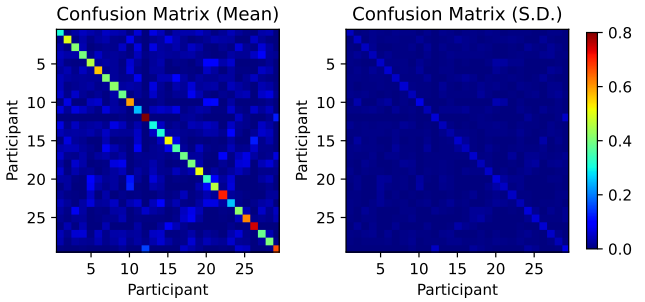

Figure 8: The above images show the cell-wise mean and standard deviation of 30 confusion matrices from PartNet.

have a single participant with a high entropy (e.g., cluster 1). These clusters exhibit common behaviors in an uncommon way, clusters 8-11 exhibit common behaviors in a common way, cluster 13 exhibits uncommon behaviors in a common way, and cluster 7 exhibits relatively uncommon behaviors in an uncommon way.

Conclusion

We conducted a virtual reality experiment investigating human navigation in a small indoor environment and successfully classified participants according to their behavioral patterns using our proposed neural network, PartNet, exceeding the accuracy of ElasticNET and the RFC. Where PartNet was lacking in interpretability was compensated by our proposed analysis pipeline. The ablation study revealed feature importance and improved feature selection for learning the embedding representation, the clustering analysis revealed the separability of some participants and the commonality of others, and the embedding analysis connected the participants back to their features for validating both the ablation study and the clustering analysis. Participants were successfully clustered into groups, suggesting the existence of navigational styles that differ primarily with respect to their distance from the midline and speed. The consistency and predictability of navigation behavior across scenarios, coupled with the varying degrees of individual distinctiveness, can inform research on egress, crowd simulation, and monitoring of large public spaces and events. Overall, our method accurately captures the dynamic nature of spatial decision-making across multiple contexts.

References

Agarap, A. (2018). Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.

Alahi, A., Ramanathan, V., Goel, K., Robicquet, A., Sadeghian, A. A., Fei-Fei, L., & Savarese, S. (2017, January). Chapter 9 - Learning to Predict Human Behavior in Crowded Scenes. In V. Murino, M. Cristani, S. Shah, & S. Savarese (Eds.), *Group and Crowd Behavior for Computer Vision* (pp. 183–207). Academic Press. doi: 10.1016/B978-0-12-809276-7.00011-4

Antonenko, P. D., Toy, S., & Niederhauser, D. S. (2012, June). Using cluster analysis for data mining in educational technology research. *Educational Technology Research and Development*, *60*(3), 383–398. (122 citations (Crossref) [2024-07-26]) doi: 10.1007/s11423-012-9235-8

Atev, S., Miller, G., & Papanikolopoulos, N. P. (2010, September). Clustering of Vehicle Trajectories. *IEEE Transactions on Intelligent Transportation Systems*, *11*(3), 647–657. doi: 10.1109/TITS.2010.2048101

Bian, J., Tian, D., Tang, Y., & Tao, D. (2018, February). *A survey on trajectory clustering analysis.* arXiv. (arXiv:1802.06971 [cs])

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10), P10008.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Brusco, M. J., & Steinley, D. (2006). Clustering, seriation, and subset extraction of confusion data. *Psychological methods*, *11*(3), 271.

Burkart, N., & Huber, M. F. (2021, January). A Survey on the Explainability of Supervised Machine Learning. *Journal of Artificial Intelligence Research*, *70*, 245–317. (465 citations (Crossref) [2024-07-26]) doi: 10.1613/jair.1.12228

Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019, August). Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics*, *8*(8), 832. (799 citations (Crossref) [2024-07-26] Number: 8 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/electronics8080832

Casali, Y., Aydin, N. Y., & Comes, T. (2022, October). Machine learning for spatial analyses in urban areas: a scoping review. *Sustainable Cities and Society*, *85*, 104050. (46 citations (Crossref) [2024-07-26]) doi: 10.1016/j.scs.2022.104050

Corbetta, A., & Toschi, F. (2023, March). Physics of Human Crowds. *Annual Review of Condensed Matter Physics*, *14*(Volume 14, 2023), 311–333. (19 citations (Crossref) [2024-07-26] Publisher: Annual Reviews) doi: 10.1146/annurev-conmatphys-031620-100450

Cristani, M., Paggetti, G., Vinciarelli, A., Bazzani, L., Menegaz, G., & Murino, V. (2011). Towards computational proxemics: Inferring social relations from interpersonal distances. In (pp. 290–297). (00072 tex.art_number: 6113127) doi: 10.1109/PASSAT/SocialCom.2011.32

Crociani, L., Vizzari, G., & Bandini, S. (2018). Between Avoidance and Imitation: Plausible Wayfinding in Pedestrian Agent-Based Models. In *Proceedings of the 19th Workshop "From Objects to Agents" (WOA 2018).* Palermo.

Dalton, R. C., Hölscher, C., & Montello, D. R. (2019). Wayfinding as a Social Activity. *Frontiers in Psychology*, *10*. doi: 10.3389/fpsyg.2019.00142

Daza, M., Barrios-Aranibar, D., Diaz-Amado, J., Cardinale, Y., & Vilasboas, J. (2021, February). An Approach of Social Navigation Based on Proxemics for Crowded Environments of Humans and Robots. *Micromachines*, *12*(2), 193. (24 citations (Crossref) [2024-07-26] Number: 2 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/mi12020193

Dong, W., Qin, T., Yang, T., Liao, H., Liu, B., Meng, L., & Liu, Y. (2022, January). Wayfinding Behavior and Spatial Knowledge Acquisition: Are They the Same in Virtual Reality and in Real-World Environments? *Annals of the American Association of Geographers*, *112*(1), 226–246. doi: 10.1080/24694452.2021.1894088

Doshi-Velez, F., & Kim, B. (2017, March). *Towards A Rigorous Science of Interpretable Machine Learning.* arXiv. (arXiv:1702.08608 [cs, stat]) doi: 10.48550/arXiv.1702.08608

Du, K. L. (2010, January). Clustering: A neural network approach. *Neural Networks*, *23*(1), 89–107. (215 citations (Crossref) [2024-07-26]) doi: 10.1016/j.neunet.2009.08.007

Farr, A. C., Kleinschmidt, T., Yarlagadda, P., & Mengersen, K. (2012, November). Wayfinding: A simple concept, a complex process. *Transport Reviews*, *32*(6), 715–743. (99 citations (Crossref) [2024-07-26] Publisher: Routledge _eprint: https://doi.org/10.1080/01441647.2012.712555) doi: 10.1080/01441647.2012.712555

Frohnwieser, A., Hopf, R., & Oberzaucher, E. (2013). Human walking behavior: the effect of pedestrian flow and personal space invasions on walking speed and direction. *Human Ethology Bulletin*.

Gérin-Lajoie, M., Richards, C., Fung, J., & McFadyen, B. (2008). Characteristics of personal space during obstacle circumvention in physical and virtual environments. *Gait and Posture*, *27*(2), 239–247. (00102 tex.author_keywords: Human locomotor control; Obstacle avoidance; Obstructed gait; Personal space; Virtual reality) doi: 10.1016/j.gaitpost.2007.03.015

Haghani, M. (2020, September). Empirical methods in pedestrian, crowd and evacuation dynamics: Part II. Field methods and controversial topics. *Safety Science*, *129*, 104760. doi: 10.1016/j.ssci.2020.104760

Haghani, M., & Sarvi, M. (2017). Following the crowd or avoiding it? Empirical investigation of imitative behaviour in emergency escape of human crowds. *Animal Be-*

haviour, *124*, 47–56. (00000 tex.author_keywords: choice uncertainty; conformity; decision making; econometric inference; emergency evacuation; herding; human crowds; pedestrian dynamics; simulated evacuations; social influence) doi: 10.1016/j.anbehav.2016.11.024

Haghani, M., & Sarvi, M. (2018). Crowd behaviour and motion: Empirical methods. *Transportation Research Part B: Methodological*, *107*, 253–294. (00000 tex.author_keywords: Animal crowd experiments; Collective motion; Crowd disasters; Crowd dynamics; Crowd management; Crowd safety; Data collection; Decision making; Emergency evacuations; Empirical observations; Evacuation drills; Experimentation; Human crowds; Lab and field data; Laboratory experiments; Operational, tactical and strategic decision; Pedestrian crowds; Virtual-reality experiments; Walking behaviour; Wayfinding) doi: 10.1016/j.trb.2017.06.017

Hall, E. T. (1968). Proxemics. *Current Anthropology*, *9*(2-3), 83–108. (00908)

Haq, S., Hill, G., & Pramanik, A. (2005). Comparison of configurational, wayfinding and cognitive correlates in real and virtual settings. In (Vol. 2, pp. 387–405).

Hayduk, L. A. (1981). The shape of personal space: An experimental investigation. *Canadian Journal of Behavioural Science*, *13*(1), 87–93. (00000) doi: 10.1037/h0081114

Helbing, D., & Molnár, P. (1995). Social force model for pedestrian dynamics. *Physical Review E*, *51*(5), 4282–4286. (05241) doi: 10.1103/PhysRevE.51.4282

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448–456).

Kavanagh, J., Morrison, S., & Barrett, R. (2005). Coordination of head and trunk accelerations during walking. *European journal of applied physiology*, *94*, 468–475.

Kim, S., Guy, S. J., Hillesland, K., Zafar, B., Gutub, A. A.-A., & Manocha, D. (2015, May). Velocity-based modeling of physical interactions in dense crowds. *The Visual Computer*, *31*(5), 541–555. (60 citations (Crossref) [2024-07-26]) doi: 10.1007/s00371-014-0946-1

Kinateder, M., & Warren, W. H. (2016, July). Social Influence on Evacuation Behavior in Real and Virtual Environments. *Frontiers in Robotics and AI*, *3*. doi: 10.3389/frobt.2016.00043

Koumetio Tekouabou, S. C., Diop, E. B., Azmi, R., Jaligot, R., & Chenal, J. (2022, September). Reviewing the application of machine learning methods to model urban form indicators in planning decision support systems: Potential, issues and challenges. *Journal of King Saud University - Computer and Information Sciences*, *34*(8, Part B), 5943–5967. (18 citations (Crossref) [2024-07-26]) doi: 10.1016/j.jksuci.2021.08.007

Moran, C., Campbell, D. J., Campbell, T. S., Roach, P., Bourassa, L., Collins, Z., … McLane, P. (2021). Predictors of attitudes and adherence to covid-19 public health guidelines in western countries: a rapid review of the emerging literature. *Journal of Public Health*, *43*(4), 739–753.

Rios-Martinez, J., Spalanzani, A., & Laugier, C. (2015). From proxemics theory to socially-aware navigation: A survey. *International Journal of Social Robotics*, *7*(2), 137–153. (00137 tex.author_keywords: Human-aware navigation; Proxemics; Socially-aware navigation) doi: 10.1007/s12369-014-0251-1

Schaumann, D., Sohn, S. S., Usman, M., Haworth, B., & Kapadia, M. (2019). Spatiotemporal Influence and Affordance Maps for Occupant Behavior Simulation. *CAADFutures19*, 18. (00000)

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Sohn, S. S., Zhou, H., Moon, S., Yoon, S., Pavlovic, V., & Kapadia, M. (2020). Laying the Foundations of Deep Long-Term Crowd Flow Prediction. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer Vision – ECCV 2020* (pp. 711–728). Cham: Springer International Publishing. (11 citations (Crossref) [2024-07-26]) doi: 10.1007/978-3-030-58526-$6_4$2

Sutskever, I., Martens, J., Dahl, G., & Hinton, G. (2013). On the importance of initialization and momentum in deep learning. In *International conference on machine learning* (pp. 1139–1147).

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 267–288.

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, *9*(11).

Vizzari, G., Crociani, L., & Bandini, S. (2020, January). An agent-based model for plausible wayfinding in pedestrian simulation. *Engineering Applications of Artificial Intelligence*, *87*, 103241. doi: 10.1016/j.engappai.2019.103241

Watson, O. M. (2014). *Proxemic Behavior: A Cross-Cultural Study*. Walter de Gruyter GmbH & Co KG. (Google-Books-ID: fF9bDAAAQBAJ)

Yue, J., Manocha, D., & Wang, H. (2022). Human Trajectory Prediction via Neural Social Physics. In S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Eds.), *Computer Vision – ECCV 2022* (pp. 376–394). Cham: Springer Nature Switzerland. (23 citations (Crossref) [2024-07-26]) doi: 10.1007/978-3-031-19830-$4_2$2

Zheng, X., Zhong, T., & Liu, M. (2009, March). Modeling crowd evacuation of a building based on seven methodological approaches. *Building and Environment*, *44*(3), 437–445. (00000) doi: 10.1016/j.buildenv.2008.04.002

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301–320.