# Multimodal Trajectory Analysis of Visitor Engagement with Interactive Science Museum Exhibits

Andrew Emerson, Nathan Henderson, Wookhee Min, Jonathan Rowe, James Minogue, and James Lester

North Carolina State University, Raleigh, NC 27695, USA
{ajemerso, nlhender, wmin, jprowe, james_minogue, lester}@ncsu.edu

**Abstract.** Recent years have seen a growing interest in investigating visitor engagement in science museums with multimodal learning analytics. Visitor engagement is a multidimensional process that unfolds temporally over the course of a museum visit. In this paper, we introduce a multimodal trajectory analysis framework for modeling visitor engagement with an interactive science exhibit for environmental sustainability. We investigate trajectories of multimodal data captured during visitor interactions with the exhibit through slope-based time series analysis. Utilizing the slopes of the time series representations for each multimodal data channel, we conduct an ablation study to investigate how additional modalities lead to improved accuracy while modeling visitor engagement. We are able to enhance visitor engagement models by accounting for varying levels of visitors' science fascination, a construct integrating science interest, curiosity, and mastery goals. The results suggest that trajectory-based representations of the multimodal visitor data can serve as the foundation for visitor engagement modeling to enhance museum learning experiences.

**Keywords:** Museum Learning, Visitor Engagement, Multimodal Trajectory Analytics.

## 1 Introduction

Visitor engagement plays a critical role in museum learning [1–2]. By focusing on how to model and enhance core components of visitor engagement, museum exhibit designers can create meaningful science learning experiences that can have lasting impact beyond the original visit [3]. However, modeling museum visitor engagement poses significant challenges. Museum learning is a *free choice* experience: visitors are free to approach and leave museum exhibits at any time. This can lead to exhibit dwell times that are exceedingly short [4–6]. Leveraging multimodal trajectory analyses using computational models that incorporate temporal data holds considerable promise for measuring and predicting visitor engagement. Previous work has utilized multimodal data for tasks such as modeling learner knowledge [7–8] and engagement [9–10], but comparatively little work has explored the use of multimodal trajectories to inform models of visitor engagement in museums. It is important to develop computational models of visitor engagement that account for inherent differences in visitor characteristics and to understand which modalities provide the most predictive value to these models.

In this paper, we introduce a multimodal learning analytics framework that induces computational models of visitor engagement using multimodal trajectory analysis. We focus on modeling visitor dwell time, a measure of behavioral engagement, with an interactive exhibit about environmental sustainability, FUTURE WORLDS. The exhibit was instrumented with multiple sensors to capture visitor behavior at the exhibit, including their posture, facial expressions, exhibit interaction logs, and eye gaze. We construct a random effects linear model of visitor dwell time that accounts for different levels of visitors' *science fascination*, a construct that integrates science interest, curiosity, and mastery goals [11]. Leveraging temporal features extracted from each of the multimodal data channels, we investigate relationships among the modalities, visitors' science fascination levels, and dwell times.

## 2 Multimodal Trajectory Models in FUTURE WORLDS

FUTURE WORLDS is an interactive science museum exhibit that combines game-based learning and interactive tabletop technologies to support visitor explorations of environmental sustainability (Fig. 1). The exhibit enables learners to explore environmental sustainability problem scenarios by investigating the impacts of alternate environmental decisions on a 3D simulated environment [12]. The science content in FUTURE WORLDS is designed for learners aged 10-11.

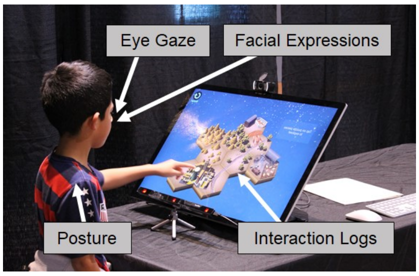


**Fig. 1.** Visitor interacting with the instrumented FUTURE WORLDS exhibit.

Multimodal visitor trajectory data was collected during a study with FUTURE WORLDS at a science museum [13]. There were 116 elementary school students aged 10-11 (*M*=10.4, *SD*=0.57) that participated in the study. The students completed a series of questionnaires and surveys before and after interacting with the exhibit. These included a demographics survey, sustainability content knowledge assessment, and the Fascination in Science scale [14] prior to exploring FUTURE WORLDS, and a sustainability content knowledge assessment and engagement survey afterward. Several participants were missing data, leaving 70 participants' data that were used for modeling. Each visitor interacted with FUTURE WORLDS individually until they successfully solved the exhibit's environmental problem or up to a maximum of approximately 12 minutes (*M*=4.13, *SD*=2.38).

This analysis specifically focuses on differences between visitors with low and high scores on the Fascination in Science scale [14]. The survey consists of eight 4-point Likert scale items. Responses are averaged to produce an overall science fascination score (*M*=3.23, *SD*=0.58). We split the 70 visitors into two groups (low and high) based on the median value of their overall science fascination scores.

**Multimodal Features.** FUTURE WORLDS was fully instrumented to track visitor behavior in real-time as described in [13]. To generate temporal representations of each visitor's interaction trajectory across their time spent exploring the exhibit, several features were generated by averaging or summing over the captured data from the start of the session to the current timestamp, using 10-second time intervals. Two posture-based features were distilled from four vertices tracked by the Microsoft Kinect, resulting in eight features. The facial expression data was used to generate 17 distinct features. Eight features were distilled from the interaction log data. Four gaze-based features were distilled to reflect the total duration of time spent fixating on specific areas of interest shown on the FUTURE WORLDS exhibit's display [13].

The distilled features were then converted into slope-based representations. Specifically, we fit an ordinary least squares (OLS) linear regression model to the series of points associated with each feature per visitor (i.e., one regression line per feature per visitor). Based upon these models, we derived a slope coefficient summarizing the trajectory of that feature for that visitor. For example, if a visitor interacted with FUTURE WORLDS for two minutes, they would have 12 data points (one data point per 10-second interval) per feature per modality. Upon fitting the OLS to the 12 points for each individual feature, we use the regression model's slope to represent the feature's temporal trajectory, thus capturing the rate at which the feature changes over time.

**Multimodal Trajectory Models of Visitor Engagement.** We induce predictive models of visitor dwell time using two linear models: (1) a baseline linear regression model that groups all museum visitors into a single group, and (2) a multilevel linear regression model that introduces random intercepts dependent on visitors' science fascination levels (low or high). We also conducted ablation analyses to investigate how the accuracy of the models of visitor dwell time vary with different sets of modalities, where we start with the full set of modalities: posture (P), facial expression (F), interaction logs (I), and eye gaze (E).

**Predictive Performance.** Each of the predictive linear models was trained and evaluated using visitor-level leave-one-out cross-validation. We report the $R^2$ for each type of model and ablation condition. Additionally, we report model performance across all visitors as well as for the high (HF) and low (LF) science fascination groups.

Within each predictive linear model, the features used were the slopes of the set of multimodal features for each ablation condition. We performed feature selection within each training fold of cross-validation by conducting a univariate linear regression test between the training set features and the target variable using a significance threshold of 0.3. In addition to feature selection, all input data were standardized within the training fold by using each individual feature's mean and standard deviation.

Table 1 shows the performance of each model predicting visitor dwell time in seconds. The random effects (RE) linear model outperforms the baseline linear model in three of the four ablation conditions. This suggests that differentiating between visitors with high and low science fascination scores enables linear models to better explain the variance in dwell time under different multimodal ablation conditions. RE achieves high predictive performance for both the PFI and PFIE ablation conditions, with the PFI condition outperforming all other conditions for the entire visitor population. The ablation analyses suggest that slope-based representations of facial expression and posture data provide limited benefit; therefore it may be helpful to investigate a hybrid approach that combines extracted features and slope-based representations for those two modalities.

**Table 1.** Predictive performance of the Baseline Model and Random Effects Linear Model. Bold values indicate the best performance for a specific ablation condition and group.

| Model Type | Group | PFIE $R^2$ | PFI $R^2$ | PF $R^2$ | P $R^2$ |
|---|---|---|---|---|---|
| Baseline Model | *All Visitors* | 0.464 | 0.614 | -0.349 | **-0.082** |
| | *LF* | **0.623** | **0.612** | -0.135 | **-0.167** |
| | *HF* | 0.359 | 0.604 | -0.510 | **-0.067** |
| Random Effects Linear Model | *All Visitors* | **0.586** | **0.634** | **-0.090** | -0.106 |
| | *LF* | 0.580 | 0.602 | **-0.003** | -0.236 |
| | *HF* | **0.577** | **0.641** | **-0.171** | -0.068 |

## 3 Conclusion and Future Work

Engagement plays a critical role in museum-based learning. Modeling visitor engagement in science museums presents significant challenges, as visitor interaction with exhibits is often brief, and visitor engagement dynamics are affected by a wide range of factors. We have introduced a multimodal trajectory analysis framework for modeling visitor dwell time with interactive science museum exhibits. Results show random effects models that account for visitors' fascination in science yield more accurate models of visitor dwell time than baseline linear models. In addition, multimodal models incorporating visitor posture, facial expressions, and interaction logs outperform models with other modality configurations. Trajectory-based feature representations effectively incorporate temporal attributes of behavioral cues for modeling visitor engagement. Potential avenues for future work include investigating more temporal features within a single visitor's interactions, exploring additional factors that influence visitor engagement, and incorporating visitor engagement models into exhibits to operate at run-time to adaptively enhance visitors' learning experiences.

**Acknowledgements.**
The authors would like to thank the staff and visitors of the North Carolina Museum of Natural Sciences. This research was supported by the National Science Foundation under Grant DRL-1713545. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References


1. Hein, G.: Learning science in informal environments: People, places, and pursuits. Museums and Social Issues 4(1), 113-124 (2009).
2. Falk, J., Dierking, L.: Learning from museums. Rowman & Littlefield (2018).
3. Allen, S.: Designs for learning: Studying science museum exhibits that do more than entertain. Science education 88(S1), S17-S33 (2004).
4. Diamond, J., Horn, M., Uttal, D.: Practical evaluation guide: Tools for museums and other informal educational settings. Rowman & Littlefield (2016).
5. Lane, H., Noren, D., Auerbach, D., Birch, M., Swartout, W.: Intelligent tutoring goes to the museum in the big city: A pedagogical agent for informal science education. In: Proceedings of the 15th International Conference on Artificial Intelligence in Education, pp. 155-162. Springer, Berlin, Heidelberg (2011).
6. Long, D., McKlin, T., Weisling, A., Martin, W., Guthrie, H., Magerko, B.: Trajectories of physical engagement and expression in a co-creative museum installation. In: Proceedings of the 12th Annual ACM Conference on Creativity and Cognition, pp. 246-257 (2019).
7. Sharma, K., Papamitsiou, Z., Giannakos, M: Building pipelines for educational data using AI and multimodal analytics: A “grey-box” approach. British Journal of Educational Technology 50(6), 3004-3031 (2019).
8. Sharma, K., Papamitsiou, Z., Olsen, J., Giannakos, M.: Predicting learners' effortful behaviour in adaptive assessment using multimodal data. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, pp. 480-489 (2020).
9. Bosch, N., D’Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., Wang, L., Zhao, W.: Detecting student emotions in computer-enabled classrooms. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence, pp. 4125–4129 (2016).
10. Dhall, A., Kaur, A., Goecke, R., Gedeon, T.: Emotiw 2018: Audio-video, student engagement and group-level affect prediction. In: Proceedings of the 20th ACM International Conference on Multimodal Interaction, pp. 653-656 (2018).
11. Bonnette, R. N., Crowley, K., Schunn, C. D.: Falling in love and staying in love with science: Ongoing informal science experiences support fascination for all children. International Journal of Science Education, 41(12), 1626–1643 (2019).
12. Rowe, J., Lobene, E., Mott, B., Lester, J.: Play in the museum: Design and development of a game-based learning exhibit for informal science education. International Journal of Gaming and Computer-Mediated Simulations 9(3), 96-113 (2017).
13. Emerson, A., Henderson, N., Rowe, J., Min, W., Lee, S., Minogue, J., Lester, J.: Early prediction of visitor engagement in science museums with multimodal learning analytics. In: Proceedings of the Twenty-Second International Conference on Multimodal Interaction, pp. 107-116 (2020).
14. Chung, J., Cannady, M. A., Schunn, C., Dorph, R., Bathgate, M.: Measures Technical Brief: Fascination in Science. Retrieved from: http://www.activationlab.org/wp-content/uploads/2016/02/Fascination-Report-3.2-20160331.pdf, last accessed 2016