Joint Choice Time: A Metric for Better Understanding Collaboration in Interactive Museum Exhibits

Matthew Berland
mberland@wisc.edu
University of Wisconsin-Madison
Madison, Wisconsin, USA

Vishesh Kumar
vishesh.kumar@northwestern.edu
Northwestern University
Evanston, USA

ABSTRACT

In this paper, we propose a new metric – Joint Choice Time (JCT) – to measure how and when visitors are collaborating around an interactive museum exhibit. This extends dwell time, one of the most commonly used metrics for museum engagement – which tends to be individual, and sacrifices insight into activity and learning details for measurement simplicity. We provide an exemplar of measuring JCT using a common "diversity metric" for collaborative choices and potential outcomes. We provide an implementable description of the metric, results from using the metric with our own data, and potential implications for designing museum exhibits and easily measuring social engagement. Here, we apply JCT to an interactive exhibit game called "Rainbow Agents" where museum visitors can play independently or work together to tend to a virtual garden using computer science concepts. Our data showed that diversity of meaningful choices positively correlated with both dwell time and diversity of positive and creative outcomes. JCT - as a productive as well as easy to access measure of social work - provides an example for learning analytics practitioners and researchers (especially in museums) to consider centering social engagement and work as a rich space for easily assessing effective learning experiences for museum visitors.

CCS CONCEPTS

• **Applied computing** → *Interactive learning environments*; **Collaborative learning**; • **Human-centered computing** → **Empirical studies in HCI**.

KEYWORDS

Museums, Collaboration, Games, Social participation, Engagement, Play

ACM Reference Format:
Matthew Berland and Vishesh Kumar. 2023. Joint Choice Time: A Metric for Better Understanding Collaboration in Interactive Museum Exhibits. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3576050.3576088

LAK '23, March 13–17, 2023, Arlington, TX, USA

ACM ISBN 978-1-4503-9865-7/23/03…$15.00
https://doi.org/10.1145/3576050.3576088

1 INTRODUCTION

Science museums are envisioned as spaces for people to explore different phenomena in ways that they can see, experiment with, and sense-make. Museums are uniquely situated in the educational ecosystem in that they are explicitly designed to support many simultaneous "social configurations" in the visitorship [6]. People come alone, with family, friends, school groups etc. The social potential of museums – to interact in novel ways with acquaintances as well as strangers – is often considered central to the experience by designers [10], but there are few quantitative measures of how those social configurations affect how people learn in the museum. The question is not new – Loomis[10] identified the need for research and design around how people collaborate to learn in museums. Recent work [6, 18] builds on Loomis by identifying how visitors engage simultaneously with both exhibits and other museum visitors in ways that the exhibit or interactive did not intend. This may be because those exhibits are not designed for collaboration (or even social interaction) but instead place "the individual's interaction with the artefact or system at the heart of the agenda"[6]. This agenda and the resulting designs can effectively split museum visitors' attention, hampering both the learning intended by the design and the potential benefits of social learning. Designing exhibits to specifically encourage pathways of social interaction and learning has the potential to surface the social and community nature of the learning experience[6].

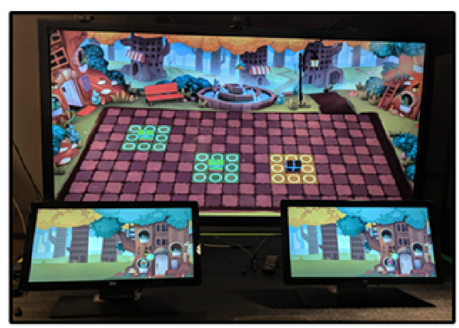

Figure 1: Rainbow Agents on the exhibit floor

Accordingly, the designers sought to foreground social interaction around computer science content in designing the Rainbow Agents exhibit game [14], a computing education game which situates players as programmers of animals who sow seeds and water plants in a virtual garden (see Figure 1, above). Visitors program their Rainbow Agents to successfully tend to and grow a flourishing and vibrant garden – using different programmable state machines – sequential, probabilistic, and stochastic – presented as friendly garden animals (a bird, a hedgehog, and a salamander). As in Figure 1, two touchscreen-based controllers are situated in front of a large, shared display, drawing on earlier works around learning through shared performances[11]; a visitor may be a player, an observer, and an advisor when the space is designed in this way. As described further in the next section, this game involves three kinds of decision making – choosing a programmable *agent* (one of the animals); placing them on the garden at an available spot; and programming them using a simple programming interface and instructional cards (which include the ability to plant one of 3 kinds of seeds in one of the 8 neighboring locations around where an agent is placed, or watering any one of the 8 spots). This "programming" interacts with randomly appearing treasure boxes on the garden that reward players with special plants they can collaboratively place on the garden, and also contributes to the garden's biodiversity which leads to a whole-garden thunderstorm when it reaches certain thresholds. The design of this game is primarily aimed at middle school youth, which appropriately corresponds to the dominant demographic of visitors at the museums this game is placed in. At the same time, our data only looks looks at gameplay logs with no off-screen information about the visitors, so it includes the whole spectrum of extremely young visitors (pres-school and elementary youth) to adult visitors who also interact with the game. We believe that this diversity does not particularly affect the phenomenon of interest in our study, since different kinds of learning and social play are even fostered in more diverse ways including visitors across age gaps as well.

One broader goal is to better understand how to develop experiences and games that support collaborative learning across social configurations. Toward that end, we present an analysis of Rainbow Agents gameplay where we try to understand the efficacy of joint gameplay (often enriched through off-table interactions) through visitors' in-game actions. This is not new to this paper, as using learning analytics to understand collaborative museum experiences has been growing[12]. Analytics about individual digital museum exhibits has explored similar phenomena: examining how people explore different challenges[9], how they socially construct rich simulations [13], and how their work can reflect engaging in productive exploration or unproductive struggle[21], among other questions.

Despite this line of work, museum practitioners tend to prioritize simplicity and accessibility of metrics for assessing visitor engagement at different exhibits — most commonly seen in the form of dwell time [8]. The length of time spent at an exhibit is used to decide if an exhibit is successful in terms of visitor engagement and assumed learning. We recognize the value of centering this simplicity in design and implementability – practitioners can measure success and engagement at all kinds of exhibits, simply and quickly without extensive reconsideration. Our design of Joint Choice Time prioritizes this conceptual and implementation simplicity, such that it can be adapted across different museum exhibits as simply as possible. In this work, we combine the description and measure of Joint Choice Time work at Rainbow Agents with an exhibit specific measure of in-game actions that represent successful play and learning, and test whether Joint Choice Time is effective as a measure of deeper engagement and learning. This centering of simplicity and accessibility dictates our focus in using only gameplay log data, and not any other sources that would help provide us richer data around social (especially off-screen) interactions occurring at the exhibit. Since extraneous data sources would require an added layer of infrastructural support (installation of cameras or microphones) as well as calibration at the data analytic level, it would move us significantly away from proposing metrics that all kinds of museums can use in a variety of exhibits. This particularly invites the question whether joint play from only gameplay log data can meaningfully highlight episodes and measure events coinciding with greater success and learning. To summarize, the research questions in this work are the following:

- How can we conceptualize and measure social work at interactive museum exhibits in the form of Joint Choice Time?
- Does greater Joint Choice Time play correlate with greater in-game success and learning?

2 JOINT CHOICE TIME: A LENS ON INTERACTIVE MUSEUM EXHIBIT DESIGN

Interactive museum exhibits can be tricky to measure, in part due to a diffuse set of learning goals, modalities, and perspectives. In this paper, we detail a new metric – Joint Choice Time (JCT) – designed to measure how visitors interact with an exhibit and with each other in ways that prioritize making choices simultaneously. This metric is contrasted with simpler "count" metrics such as "time on task" or "joint dwell time" which may lean negative as they increase. JCT, on the other hand, prioritizes the meaningful or disciplinary decisions that visitors make in an interactive exhibit. For simplicity, we provide an example using only logs (i.e., no identifiable or qualitative data) collected from Rainbow Agents, a game designed to teach computer science. We find that as JCT increases, all in-game "scoring" metrics also increase (discussed in Results).

Having argued for the relevance of choices, specifically collaborative choices, in understanding cooperation and collaboration in a museum exhibit, there is value in developing a (hopefully simple and portable) metric for quantifying, exploring, or identifying those choices. We call this "joint choice time" and, fundamentally, it is a quantified catalog of "choice moments."

A "choice moment" comes from both education literature [17] and game design literature [3, 16]. Here, we will build our notion of the choice as a "consequence [for further gameplay that must] alter the game" from Fullerton, though they do not necessarily differ meaningfully in the context of Rainbow Agents. In Rainbow Agents, almost every user action that results in a change to the game state requires a choice, therefore, our "choice moments" are every user action that makes a change to the game state (where game state is defined as "the canonical representation of the game world"). For instance, dragging a card without setting it in a "card slot" generates visual feedback, and it is (in some sense, typically) the choice of a human – they can drag things around a screen to see what it looks like. However, since this does not change the game state, we do not register it as a "choice moment." Our choice moments fall into a few obvious categories – exhibit/physical decisions (e.g., turn on or off the exhibit), placement of agent-flags (e.g., hedgehog, lizard), choice of cards (e.g., water the three top squares), and inter-player coordination actions (in this game, a pop-up which asks players to place a rainbow mushroom). Of those four categories, two are rare:

- In our data, the exhibit was never, effectively, turned on or off – it was left on for weeks; and
- Inter-player coordination actions, which, in this case, are redundant with the "placement" action at our granularity.

Therefore, we have two categories of choices: placement and cards. Both types of actions occur in abundance in our sample, as you will see in the data below. Now that we have "choice moments," we need to answer two questions:

- Which actions are joint actions?
- Which joint/solo actions are "meaningful" to us and to them?

In our data set, we have data about which "player sessions" happened on one touchscreen and which player sessions happened on two touchscreens. Since the screens are near each other, and the viewscreen reflects the actions of both touchscreens, we suggest that all actions in which two people are acting at the same time are, effectively, joint actions. They may not be deeply cooperative or collaborative, but they are, at very least, joint. It is very rare in our data (we could not find an obvious example, though it must have happened) that multiple people are attending to the same touchscreen when nobody is on the second touchscreen; therefore, the one-touchscreen actions are solo. We have also never seen the other possibility of one person operating the two touchscreens in rapid consecution. The screens are far away enough that actively operating them together is significantly challenging and impractical, and never having seen it happen in our analysis of the video and our presence at the exhibit enable us to confidently assert that logs indicating actions on both screens are likely coming from different players. The EXTIRE framework [1] suggests that exploration of and tinkering with both choices and outcomes can be a useful proxy for understanding creative, playful learning. As such, we used the standard Simpson's Diversity Index (SDI) [19], a simple metric applied widely across fields [2, 20], to see the range of exploration both of choices and outcomes in Rainbow Agents. Though the fit is not perfect – SDI "prefers" a wider variety of possibilities with a smaller number of "data collection events" – it is a well-tested metric and has also been used in other visitor-focused museum education research (e.g., Roberts et al. [15]). In this case, we are only looking for "SDI per session," as it is not particularly instructive to see how it grows over time spans that typically stay under 20 minutes. Additionally, analyzing SDI across sessions would only make sense if the sequence of players and sessions influenced each other meaningfully or built on each other's work, which is rarely the case in Rainbow Agents. On the contrary, the game resets most player actions and plants fade away in a few minutes of inaction. The basic SDI measure represents the probability that any two randomly chosen entities are of the same type. So, for a sequence of actions (in terms of different machines used, plants planted, and orbs filled in gameplay up to a given point of time), more exploratory behavior would be represented by a lower probability that randomly chosen pairs of actions look the same. In the language of SDI, this would be described as the odds of two randomly chosen organisms belonging to the same species. The mathematical definition of the SDI (as per [19]) is the sum across the proportional abundance of different 'species' in the dataset. For small samples, we assume sampling without replacement and for easier representation of the value (higher index representing higher diversity), we use the Gini-Simpson variation [7], which leads to the metric used in our work as:

$$\lambda = 1 - \frac{\sum_{i=1}^{R} n_i \ (n_i - 1)}{N(N-1)} \tag{1}$$

where n_i is the number of organisms of the i^{th} species, R is the total number of species of interest, and N is the total number of organisms. This metric can be calculated over time buckets (i.e., all actions per predefined time span), action buckets (i.e., number of actions in a bucket) or in relation to other kinds of diversity (i.e., how does the diversity of the plants that have been planted and different orbs filled relate to diversity of cards used). Most critically, to recognize the diversity that is created because of collaborative work, we will contrast changes in diversity over time when visitors are working by themselves with patterns during sessions when they are working together (i.e., both controllers are being used). In our analyses, *SDI outcomes* is the SDI of the different plant possibilities on the rainbow agents shared garden; *SDI choices* is the SDI of the different card and machine possibilities across players. *Minutes played* is the number of minutes in either a joint or solo session with the exhibit. *Active players* is 1 if someone was using the exhibit alone, and 2 if two people were playing at the same time (i.e., jointly).

3 RESULTS

The definitions outlined above articulate the answer to our first research question – how can we conceptualize social play and work at an interactive exhibit in the form of Joint Choice Time. Therein, we describe how the Simpsons Diversity Index (SDI) for aspects of the game can be an effective measure of productive gameplay, especially since it describes whole-game measurement at a level that can well describe individual progress as well as group work. To establish whether engaging in joint play does indeed lead to more in-game progress, we compare the SDI metrics for when there is just one player at the game (i.e. only one screen is being played on) or 2 (or more) players are engaging across both screens.

Table 1 presents the statistics comparing gameplay and success metrics across our data for 1 player (i.e. solo play) and 2 players playing simultaneously (joint play). Given the museum nature of our exhibit, we see over 4x of the visitors engaged in joint play (1373 vs. 318), but this also coincides with an over 3x difference of average play time (2.93 minutes vs 9.11 minutes). We see that the difference of SDI outcomes in our game – an overly simplistic representation of game progress in the form of different plants planted – is also greater by a factor of almost 1.5x (0.74 vs 0.54), and the average SDI choices – the diversity of programming cards used by players – is over 3x greater (0.36 vs 0.11). The logistic regression (Table 2) indicates how there is a significant influence of both the number of joint choices and the total number of choices on the diversity of the outcomes. This comparison estimates how the change of time played is likely not the only factor leading to the increased SDI values for joint play in comparison to solo play.

Table 1: Factors by number of active players

	Players	n	mean	std
Minutes Played	1	318	2.93	4.72
	2	1373	9.11	10.61
SDI Outcomes	1	318	0.54	0.40
	2	1373	0.74	0.28
SDI Chioces	1	318	0.11	0.24
	2	1373	0.36	0.33

Table 2: Regression predicting sdi of outcomes from joint factors

	coeff	stderror	t	P > |t|
Intercept	0.3414	0.028	12.189	0.000
minutes played	0.038	0.001	4.191	0.000
active players	0.1520	0.016	9.403	0.000
sdi choices	0.1736	0.028	6.241	0.000

4 DISCUSSION AND CONCLUSIONS

The metrics of JCT not only help recognizing the efficacy of joint play and measuring it through the lens of in-game choice moments that act as markers of increasing action *diversity*, but also provide a valuable step towards identifying moments in which people are having meaningful interactions with the exhibit and with each other. That is, creating more game-specific metrics to mark increases in SDI metrics, and creating models to measure these states Joint Choice Time (JCT) is a relative measure of the diversity of collaboratively making exhibit-impactful choices and the diversity of possible outcomes. Both diversity measures come from the standard "Simpson's Diversity Index" (originally from ethology) in ways that are applied to game log data. While the use of "diversity of experience" as a proxy for valuable interactions or valuable outcomes may at first seem counter-intuitive, this approach is grounded both in theories of learning [17] and in theories of game design, especially educational game design [4]. JCT is simple to code and trivial to compute. Finally, one can readily design systems to maximize it. Our results consistently show two things:

- Visitors who work with other people spend more time making choices (visible in the 3x average time spent of 9.11 minutes vs. 2.93 minutes).
- As visitors spend more time making choices together, more positive outcomes result (regardless of time spent) (evident in the 0.1736 coefficient of SDI choices' impact on SDI outcomes).

Furthermore, unlike more commonly used museum outcome measures, JCT is less amenable to being "gamed." In contrast, when dwell time is gamed – e.g., making the same task take longer – that may feel to visitors like a net negative; it is a waste of time and often takes away from interactions of educational value [5]. However, JCT as a measure motivates both collaborative choice making and diversity of outcome. If visitors are voluntarily spending time making a diverse array of choices with other visitors in ways that generate a diverse array of outcomes, it is hard (though not impossible) to imagine how that would not feel like a meaningful, social, educational experience for both parties. At very least, the diversity of both choice and outcome is consistent with best practices in game design in ways that most measures may not be [3].

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1713439.

REFERENCES

[1] Matthew Berland, Taylor Martin, Tom Benton, Carmen Petrick Smith, and Don Davis. 2013. Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences* 22, 4 (2013), 564–599. Publisher: Taylor & Francis.

[2] P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin, and S. Zeger. 2009. *Springer series in statistics.* Springer.

[3] Tracy Fullerton. 2014. *Game design workshop: a playcentric approach to creating innovative games.* CRC press.

[4] James Paul Gee. 2004. Learning by design: Games as learning machines. *Interactive educational multimedia: IEM* (2004), 15–23.

[5] Tony Hall and Liam Bannon. 2006. Designing ubiquitous computing to enhance children's learning in museums. *Journal of Computer Assisted Learning* 22, 4 (2006), 231–243.

[6] Christian Heath and Dirk Vom Lehn. 2013. Interactivity and Collaboration: new forms of participation in museums, galleries and science centres. In *Museums in a digital age.* Routledge, 266–280.

[7] Lou Jost. 2006. Entropy and diversity. *Oikos* 113, 2 (2006), 363–375. Publisher: Wiley Online Library.

[8] Joel Lanir, Tsvi Kuflik, Julia Sheidin, Nisan Yavin, Kate Leiderman, and Michael Segal. 2017. Visualizing museum visitors' behavior: Where do they go and what do they do there? *Personal and Ubiquitous Computing* 21, 2 (2017), 313–326.

[9] Duri Long, Tom McKlin, Anna Weisling, William Martin, Hannah Guthrie, and Brian Magerko. 2019. Trajectories of physical engagement and expression in a co-creative museum installation. In *Proceedings of the 2019 on Creativity and Cognition.* 246–257.

[10] Ross J. Loomis. 1974. Social Learning Potentials of Museums. (1974). Publisher: ERIC.

[11] Leilah Lyons, Michael Tissenbaum, Matthew Berland, Rebecca Eydt, Lauren Wielgus, and Adam Mechtley. 2015. Designing visible engineering: supporting tinkering performances in museums. In *Proceedings of the 14th International Conference on Interaction Design and Children.* 49–58.

[12] Aditi Mallavarapu, Leilah Lyons, and Stephen Uzzo. 2022. Exploring the utility of social-network-derived collaborative opportunity temperature readings for informing design and research of large-group immersive learning environments. *Journal of Learning Analytics* 9, 1 (2022), 53–76.

[13] Aditi Mallavarapu, Leilah Lyons, Stephen Uzzo, Wren Thompson, Rinat Levy-Cohen, and Brian Slattery. 2019. Connect-to-connected worlds: Piloting a mobile, data-driven reflection tool for an open-ended simulation at a museum. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–14.

[14] Anthony Pellicone, Leilah Lyons, Vishesh Kumar, Eda Zhang, and Matthew Berland. 2019. Rainbow Agents: A collaborative game for computational literacy. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts.* 597–604.

[15] Jessica Roberts, Leilah Lyons, Francesco Cafaro, and Rebecca Eydt. 2014. Interpreting data from within: Supporting humandata interaction in museum exhibits through perspective taking. In *Proceedings of the 2014 conference on Interaction design and children.* 7–16.

[16] Jesse Schell. 2008. *The Art of Game Design: A book of lenses.* CRC press.

[17] Daniel L. Schwartz and Dylan Arena. 2013. *Measuring what matters most: Choice-based assessments for the digital age.* The MIT Press.

[18] Neta Shaby, Orit Ben-Zvi Assaraf, and Tali Tal. 2019. An examination of the interactions between museum educators and students on a school visit to science museum. *Journal of Research in Science Teaching* 56, 2 (2019), 211–239. Publisher: Wiley Online Library.

[19] Edward H. Simpson. 1949. Measurement of diversity. *Nature* 163, 4148 (1949), 688–688. Publisher: Nature Publishing Group.

[20] Thomas Richard Edmund Southwood and Peter A. Henderson. 2009. *Ecological methods.* John Wiley & Sons.

[21] Mike Tissenbaum, Vishesh Kumar, and Matthew Berland. 2016. Modeling Visitor Behavior in a Game-Based Engineering Museum Exhibit with Hidden Markov Models. *International Educational Data Mining Society* (2016). Publisher: ERIC.