Understanding Conceptual Engagement and Accuracy in an Assessment Game

Allison Mishkin, BrainPOP
Will Jordan-Cooley, BrainPOP
Kevin Miklasz, BrainPOP

Abstract: We are conducting an exploratory analysis into the clickstream data produced by a new "assessment game" appearing on BrainPOP, the popular educational resource. The game involves ordering and sorting information along a *Timeline*, and easily supports plug and play content from different subject areas. In this poster, we will present preliminary data on this game's performance as an assessment tool. We will analyze clickstream data to characterize common user patterns, map those user data onto a framework for conceptual accuracy and engagement, and test the framework's validity.

Introduction

Games have great potential as assessment tools, particularly for performance-based assessment (Steinkuehler and Squire, 2013). Digital games can automatically and unobtrusively observe student performance without the time-intensiveness of video recording or observational work. Clickstream data may systematize assessment if a student's sequence of clicks can be tied to learning claims, such as through evidence-centered design (Shute and Ventura, 2013). This holds great educational potential, particularly if assessment techniques can be generalized across different kinds of content or games (Owen et al. 2014). Thus, creating and validating assessment games' efficacy becomes critical when exploring the future of learning games.

An assessment game called *Time Zone X* is intended to work as both a game and an assessment tool across any given content area. In this poster, we will analyze clickstream data from students' gameplay and adopt an exploratory approach to understanding how players interact with the game and how the game itself can act as an assessment tool for student knowledge.

Time Zone X Game Description

Time Zone X is loosely based on the physical card game, *Timeline* (Frederik, 2015). Players have cards with chronological events on them and they must accurately place those events in a sequence (Figure 1). Events are grouped based on BrainPOP movie topics into separate "decks" and players decide which deck to play from each turn (Figure 1). Completing a deck in a single game session awards the player a historical artifact (relevant to game theming). The goal of the game is to create a *Timeline* that is as long as possible. Placing multiple correct events in a row lets the player unlock new decks and extend the length of their *Timeline* with that deck's cards (Figure 1). Placing multiple incorrect events in a row ends the game.

The game, featured on BrainPOP's GameUP platform, is referred to as an "assessment game." It was designed specifically to assess student knowledge on BrainPOP topics. Assessment games can incorporate content from many different topic areas, but all rely on a single assessment framework. Additionally, this particular framework (described in the next section) is intended to be flexible enough to apply across different assessment games. The framework has only been implemented on one previous game, so this will be the first attempt to apply the framework across games.

The Analysis

The analysis will be supported through repeated user testing sessions. These sessions will help us explore user play patterns and refine the assessment framework.

Before each session, we will work with a teacher to understand what content his or her class is expected to know. We will then set up two versions of *Time Zone X*: one with content that students have recently reviewed and one with content that students are unlikely to have seen before (i.e. content targeted one or two grade levels above the students). The class will be separated into two groups. Each group will play one version of the game and then switch to the other halfway through the class. We expect to run this experiment at 2-3 user testing sessions, gaining gameplay data from about 60 students.

An exploratory analysis will be conducted on this data to identify dominant patterns of user interaction. For each student we will have two sets of gameplay data: one expected to contain a higher content understanding and one

expected to showcase lower content understanding. We can also group the data by first and second play-through of the game. This allows us to simultaneously test for the effects of prior content understanding and game familiarity.

Building on the engagement data, we will apply BrainPOP's existing assessment framework to *Time Zone X*. In a previous game called *Sortify* (BrainPOP, 2013), a two-pronged approach was developed to report on *conceptual accuracy* and *engagement* from student gameplay. In this framework, conceptual accuracy is defined as percent correct out of total attempted items (as related to a particular concept in a given activity). Meanwhile, conceptual engagement refers to *how much* a player chose to use a concept that was available to them in a given activity.

We have applied these general definitions of concept-specific accuracy and engagement to *Time Zone X* gameplay. In *Time Zone X*, we interpret accuracy as number of correctly placed events out of total attempted events. Incorrect answers will be weighted by the distance from the correct time interval. Meanwhile, we define engagement, in this context, as the total number of events related to a certain concept that were used by a player. A small qualification: the exact application of this framework to *Time Zone X* will be adjusted in ways that seem reasonable, based on the results of the exploratory analysis.

This presentation will involve two analyses. First, we will present an exploratory analysis of the clickstream data, highlighting user interaction patterns. Second, we will determine if the definitions of conceptual accuracy and engagement are correlated with prior content understanding or familiarity with the game mechanics. This will help us determine if the BrainPOP framework can generalize across different content and assessment games.

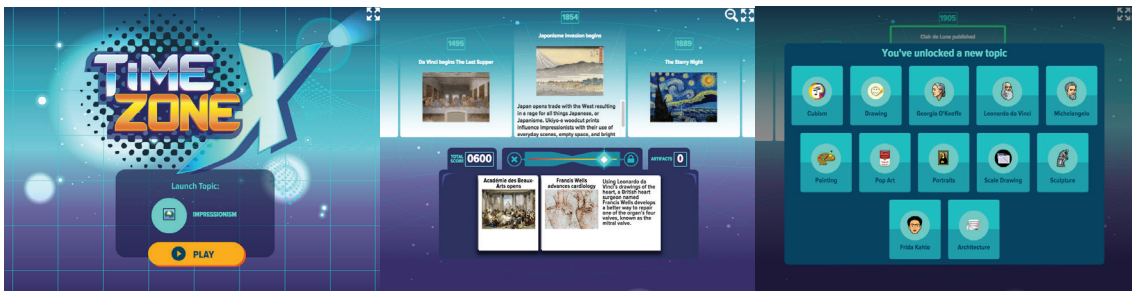

Figure 1: On the left, a screenshot of the starting screen, which shows the starting (titles directly correspond to BrainPOP topics). In the middle is a typical image from gameplay. The bottom portion of the screen holds the decks that a player can choose from while the top depicts the current *Timeline*. The meter in the middle of the screen shows the current correct (to the right) or incorrect (to the left) placement streak. Currently, the meter is one correct placement away from unlocking a new deck. The right screen shows how new decks are selected, after being unlocked by a series of correct placements.

References

BrainPOP. (2013). *Sortify*. https://www.brainpop.com/games/sortify/. New York: BrainPOP.

Henry, Frederik. (2015). *Timeline*. http://us.asmodee.com/ressources/jeux_versions/*Timeline* _3.php. Montreal: Asmodee.

NGSS Lead States. (2013). Next Generation Science Standards: For States, By States. Washington, DC: The National Academies Press.

Owen, V.E., Ramirez, D., Salmon, A. & Halverson, R. (2014). Capturing Learner Trajectories in Educational Games through ADAGE (Assessment Data Aggregator for Game Environments): A Click-Stream Data Framework for Assessment of Learning in Play. Presented at the 2014 American Educational Research Association Annual Meeting, Philadelphia, PA.

Shute, V. J. & Ventura, M. (2013). Measuring and supporting learning in games: Stealth assessment. Cambridge, MA: The MIT Press.

Steinkuehler C. and Squire K. (2013). Videogames and Learning. *Cambridge Handbook of the Learning Sciences*, New York: Cambridge University Press.