Educational games have become an established paradigm of instructional practice; however, there is still much to be learned about how to design games to be the most beneficial for learners. An important consideration when designing an educational game is whether there is good alignment between its content goals and the instructional behaviors it makes in order to reinforce those goals. Existing methods for measuring alignment are labor intensive and use complex auditing procedures making it difficult to define and evaluate this alignment in order to guide the educational game design process. This thesis explores a way to operationalize this concept of alignment and demonstrates an analysis technique that can help educational game designers measure the alignment of both current educational game designs and prototypes of future iterations.

In my work, I explore the use of Replay Analysis, a novel technique that uses in-game replays of player sessions as a data source to support analysis. This method can be used to capture gameplay experience for the evaluation of alignment as well as other forms of analysis. The majority of my work has been performed in the context of RumbleBlocks, an educational game that teaches basic structural stability and balance concepts to young children. Using Replay Analysis, I leveraged replay data during a formative evaluation of RumbleBlocks to highlight that the game likely possesses misalignment in how it teaches some concepts of stability to players. The results led to suggestions for various design iterations.

Through exploring these design iterations, I further demonstrate an extension of Replay Analysis called Projective Replay Analysis, which uses recorded student replay data in prototypes of new versions of a game to predict whether the new version would be an improvement. I implemented two forms of projective replay: Literal Projective Replay, which uses a naïve player model that replays past player actions through a new game version exactly as they were originally recorded; and Flexible Projective Replay, which augments the process with an AI player model that uses prior player actions as training data to learn to play through a new game. To assess the validity of this method of game evaluation, I performed a new replication study of the original formative evaluation to validate whether the conclusions reached through virtual methods correspond to those reached in a normal playtesting situation. Ultimately, my findings were that Literal Projective Replay was able to predict a new and unanticipated misalignment with the game, but Flexible Projective Replay, as currently implemented, has limitations in its ability to explore new game spaces.

This work makes contributions to the fields of human-computer interaction, by exploring the benefits and limitations of different replay paradigms for the evaluation of interactive systems; learning sciences, by establishing a novel operationalization of alignment for instructional moves; and educational game design, by providing a model for using Projective Replay Analysis to guide the iterative development of an educational game.

A copy of the document can be found at: [www.erikharpstead.net/thesis](http://www.erikharpstead.net/thesis)