# – Evaluation

## Introduction

The main motivation of this chapter is to answer the following research questions defined at Chapter 1:

* Is it possible to use provenance analysis in order to understand the events transpired during the game?
* Is it faster to analyze the game by using provenance instead of watching a replay of the session?
* Does provenance analysis help to find answers with more accuracy than watching a replay of the game session?

Experiments were elaborated in order to answer these questions.

## Experiment Planning

In order to verify the possibility of using provenance to aid understanding of events in a game, it would be required to analyze a game session. Initially, it was planned to allow the volunteers to play the game and to answer a questionnaire at the end of the game session. Half the volunteers would answer the questionnaire while having access to the provenance of the game session. The other half would answer the questionnaire by only using his/her memory. However, the formulation of the questionnaire would be impractical due to the nature that each game session is different from another, even when restrictions are placed in order to control the outcome. Events that occurred to one player might not be true to another. Thus, the questionnaire would be required to cover almost all possibilities or have generic questions.

To deal with this issue, we opted for a more controlled environment. Instead of playing the game, all volunteers would watch the same game session being played by a third person. Thus, solving the problem of having multiple, and different, game scenarios. However, the problem of formulating the questionnaire would still persist. To solve this, we decided to previously record a game session and the provenance information. Then, the questionnaire would be elaborated. Thus volunteers would watch the recorded game session instead of watching it in real time. The questionnaire can be customized to the game session because it was previously recorded. This allows for specific questions to be asked about specific events and their respective outcomes in the questionnaire, instead of making generic questions trying to guess the possible outcomes.

Thus, volunteers were required watch a gameplay video and answer a questionnaire. To answer the question about using provenance analysis for aiding understanding, some volunteers answered the questionnaire by using the *Proof Viewer*, while the others only by re-watching the gameplay video.

## Experiment Execution

Due to the unfamiliarity with the game and the *Proof Viewer*, the experiment execution followed the plan detailed at the previous section with a few additions: tutorials. The execution process is divided in three stages, as illustrated by Figure 1: Generating the questionnaire, running the experiment with volunteers, and analyzing the results.

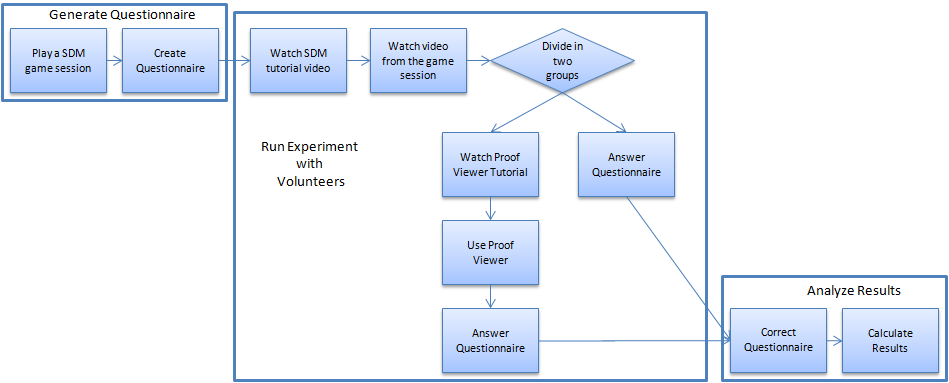


Figure : Experiment Execution Flowchart

The first stage is generating the questionnaire.

The next stage is running the experiment with volunteers. Before participating in the experiment, volunteers were required to read and sign a consent form. First, volunteers were required to watch a tutorial video about the SDM, which explained details about the game’s interface and also had a written document explaining key features. Then they were allowed to watch the gameplay video and were divided in two groups: those that would use *Proof Viewer*, and those that would not. After the gameplay video, the volunteers answered the questionnaire. However, the group for *Proof Viewer* saw a tutorial video and got a written document before being allowed to answer the questionnaire, both detailing the *Proof Viewer’s* interface and explaining all features in it. The other group was allowed only to re-watch the gameplay video. All documents used at this stage are available at Appendix A.

Lastly, analyze results.

The results obtained are: …

## Statistical Analysis

Anova.

## Threats to Validity

First time with SDM.

First time with Proof Viewer.

Few volunteers.

Questionnaire with multiple questions.

Questionnaire’s correct answers.

Time restriction.

First period students (prog1).

## Final Considerations