Abstract. The purpose of this study is to design a metric of success for an Intelligent Tutoring System (ITS) which gives programming problems for students to solve.

Description. We have partially constructed an ITS (intelligent tutoring system) which can issue problems for the student to solve. We would like to examine the differences (particularly in performance) between solving the problems through an ITS versus a conventional paper-based approach.

After we seat the student in a private and distraction-free conference room, the student will solve problems on the ITS or on paper for up to an hour, after which the system will cue the student that the trial is concluded. Once the trial has concluded, the test administrator will stop the ITS (if applicable) for the debriefing phase.

We will conduct at least one experiment, which is to test mean performance differences between the paper-based and computer-based contexts. Further experiments may test differences between two distinct problem sets given by the ITS; for example, performance differences due to differing progressions of Bloom taxonomic levels, problem difficulties, or concepts. All further experiments will test performance differences due to manipulations on the problems (rather than the interface, the test administrator's instructions, or other potential variables).