

# Collaborative Research: Frameworks: Cyber Infrastructure for Shared Algorithmic and Experimental Research in Online Learning

## Worcester Polytechnic Institute

Neil T. Heffernan (PI)

George Heineman (Co-PI)

Korinn Ostrow (Co-PI) & Anthony Botelho (Co-PI)

## University of Pennsylvania

Ryan Baker (PI)

Rebecca Stein (Co-PI)

Peter Decherney (Co-PI)

## ASSISTments

### The Platform

- ✧ Assists students via feedback and tutoring
  - Over 50,000 students around the world
  - Over 12 million problems solved last year
- ✧ Provides teachers with formative assessments
- ✧ Supports embedded randomized controlled experimentation for researchers

### An Effective Learning Tool

- ✧ Positive effects in a recent large-scale U.S. Department of Education Efficacy Trial
- ✧ Gains nearly doubled the amount students learned each year compared to expected gains on a standardized instrument
- ✧ Gains were especially impressive for underprivileged students, with the system working to close the achievement gap
- ✧ Nearly \$7 million in IES funding now supports ongoing Effectiveness Trial and Scale Up efforts

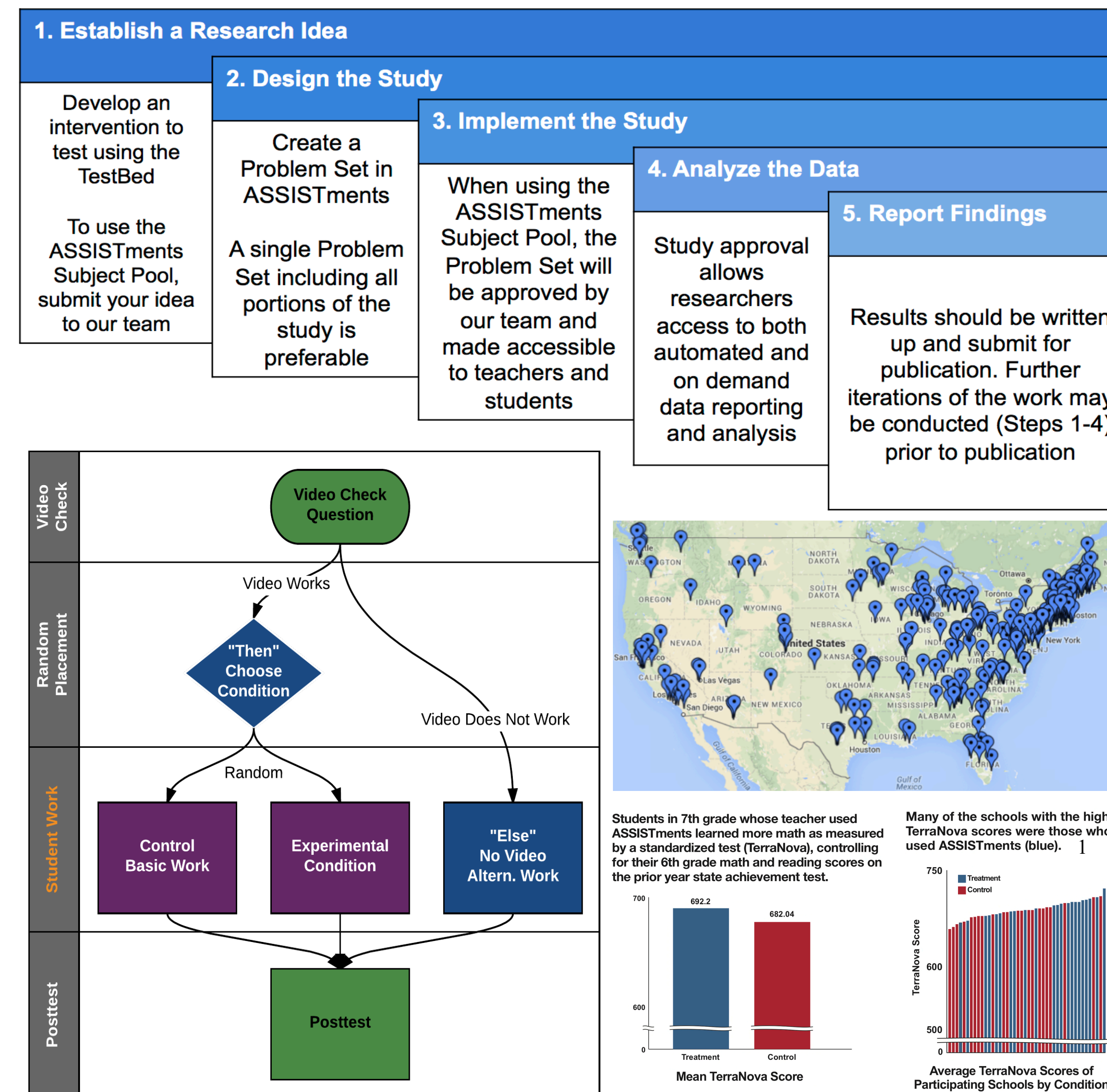
### A Successful Shared Scientific Tool

- ✧ Established in 2014, the Testbed allowed researchers from around the world to use ASSISTments for experimentation
  - Originally funded by an NSF SI2 grant
  - Leverages the ASSISTments user population
  - More than 100 studies already conducted
  - Users from top institutions and research firms
  - 5+ publications produced without ties to WPI team

### ALI: Research Reporting Infrastructure

- ✧ The Assessment of Learning Infrastructure:
  - Informal study registration process
  - Tags experimental designs for automated analysis
  - Provides access to on-demand data reports
  - Promotes open science protocols

## 2014 – 2019: The ASSISTments Testbed



Dear Researcher,

Welcome to the data record for problem set PSAMR8Z. You have received this record as part of our weekly reporting. Automated data analysis is featured below, offering a preliminary overview of your sample and a selection of analyses for your consideration. The latter portion of this report contains the raw data files from which you can conduct your own thorough analyses. When publishing your work, please reference this report as a stable location for readers to access your data for review and replication.

By downloading content from this page, you agree to our [Terms of Use](#).

### Automated Data Analysis

#### Completion Rates

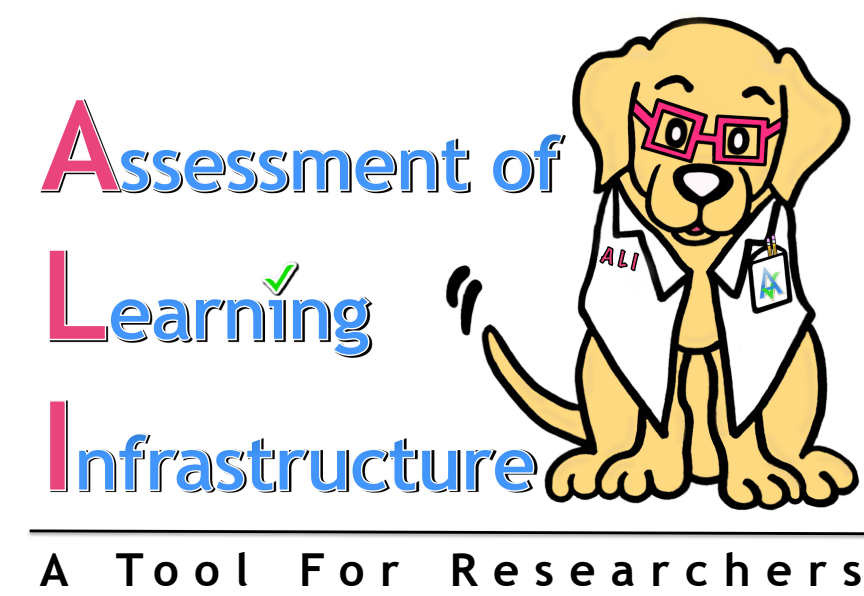
Students that have started PSAMR8Z: 1556

Students that have completed PSAMR8Z: 1043

#### Bias Assessment

Before analyzing learning outcomes, we suggest first assessing potential bias introduced by your experimental conditions (i.e., examine differential dropout). The table below reports the number of students that have completed PSAMR8Z, split out by experimental condition. Condition distribution was not significantly different,  $X^2(2, N = 762) = 4.31, p > 0.05$ .

Conditions	Started	Finished	% Completed
Text	248	144	58
Video - Human	260	148	57
Video - Pen	254	126	50



### Raw Data Files

Raw data files contain the logged information for each student that has participated in your study. We provide this data in a variety of formats, as shown below, to assist in your analytic efforts.

- *Student Covariate*
- *Action Level*
- *Problem Level*
- *Student Level*
- *Student Level + Problem Level*

For a glossary and dataset tutorials, please visit our [Glossary Page](#).

**ASSISTments**  
A Free Public Service of Worcester Polytechnic Institute

## 2020 and Beyond: E-TRIALS

### Ed-Tech Research Infrastructure to Advance Learning Science

- ✧ Recently secured funding to redesign the Testbed into E-TRIALS
  - Centralized system with one-stop researcher log-in
  - Template driven, visual canvas to aid in experimental design construction
  - Data-rich environment for easy content selection and manipulation
  - Strengthened reporting structures and automated analyses
- ✧ Support for current efforts includes:
  - \$2M – Schmidt Futures for basic infrastructure
  - \$3.3M – NSF CSSI for expansion into MOOCs
- ✧ Visit [www.etrialstestbed.org](http://www.etrialstestbed.org) to learn more today

## Goals for Current CSSI Work

### Development of RAILKAM Cyber Infrastructure

- ✧ Integration of ASSISTments' E-TRIALS with UPenn MOOCs
  - UPenn MOOCs are used by hundreds of thousands of learners each year
  - Enable broader populations, more robust student interactions, and more bountiful data collection than currently feasible in either environment alone
  - Increase the efficiency and ease of conducting quality educational research in online learning environments while targeting learning outcomes
- ✧ Project Targets
  - Support 20 researchers in conducting RCTs in K-12 + MOOC environments
  - Support 75 data scientists via redacted datasets emphasizing student privacy



Funded by the NSF (1931523, 1931419, 1724889, 1316736, 1252297, 1109483, 1031398, 0742503, 1440753), the IES (R305A120125, R305C100024), GAANN (P200A120238), and Schmidt Futures

