

---

# ASSISTments Dataset from Multiple Randomized Controlled Experiments

**Douglas Selent**

Worcester Polytechnic Institute  
100 Institute Road  
Worcester, MA 01609 USA  
dselent@wpi.edu

**Thanaporn Patikorn**

Worcester Polytechnic Institute  
100 Institute Road  
Worcester, MA 01609 USA  
tpatikorn@wpi.edu

**Neil Heffernan**

Worcester Polytechnic Institute  
100 Institute Road  
Worcester, MA 01609 USA  
nth@wpi.edu

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).  
L@S 2016, April 25-26, 2016, Edinburgh, Scotland UK  
ACM 978-1-4503-3726-7/16/04.  
<http://dx.doi.org/10.1145/2876034.2893409>

**Abstract**

In this paper, we present a dataset consisting of data generated from 22 previously and currently running randomized controlled experiments inside the ASSISTments online learning platform. This dataset provides data mining opportunities for researchers to analyze ASSISTments data in a convenient format across multiple experiments at the same time. The data preprocessing steps are explained in detail to inform researchers about how this dataset was generated. A list of column descriptions is provided to define the columns in the dataset and a set of summary statistics are presented to briefly describe the dataset.

**Author Keywords**

ASSISTments; Randomized Controlled Experiments; Dataset.

**ACM Classification Keywords**

E.0. Data: General

**Introduction***ASSISTments*

ASSISTments is a freely available online learning platform mostly used in mathematic classes between grades 4-12 in the United States. ASSISTments currently has 140,032 users who have done 32,625,908 problems from years 2013-2015. Most of

the users are middle school (grades 6-8) mathematic students, who are located in or near Massachusetts. Students use ASSISTments on classwork and homework, which may be done with or without the use of a paper copy. Instant feedback is typically provided to the students upon answering problems, so that students will immediately know if they have answered the problem correctly. Students are not allowed to skip any problems and must answer correctly in order to continue on to the next problem. If a student does not answer the problem correctly on their first attempt, the problem is marked as incorrect. Almost all problems in ASSISTments give the student the option to click on a hint button, which either offers additional help to the student or just gives the student the answer to the problem, so that he or she may continue.

#### *Experiment Infrastructure*

All of the experiments were created by either internal or external researchers working with ASSISTments. Dr. Heffernan was funded by NSF to create the ASSISTmentsTestBed.org to support external researchers to conduct studies. The ASSISTmentsTestBed.org is a reporting service that helps researchers retrieve data on experiments that have been run inside the ASSISTments system [1, 2]. This reporting service is also known as the Assessment of Learning Infrastructure (ALI), which combines automated analysis with the data reporting [6]. Several of these experiments have been previously published, with topics ranging from student confidence, student choice, and buggy messages [3-5, 8].

#### **Data Selection**

In the ASSISTments online learning platform, there are several ways to build assignment structures which experiments can be embedded inside of. One type of structure is called a Skill Builder. A Skill Builder is an assignment type that consists of a large number of similar problems, where students must answer a specified number of problems (usually three) correctly in a row on the same day in order to finish the assignment. These problems are generated by variabilized templates, where several problem instances are generated from the same template with different numbers replacing the variables of the template [7]. We chose only to include experiments that leveraged the Skill Builder infrastructure because that was the most common assignment structure. We also chose to only use experiments with a control group and a single experimental group since that was the study design for most of the experiments. Using only Skill Builder assignments with a single experimental condition allows for a common set of independent and dependent variables to be used for all 22 experiments.

To ensure that there was a large enough sample of students for each experiment, only experiments that had over fifty students complete the assignment in all the conditions of the experiment were chosen. Also, only experiments that had less than 10% corrupted data were used.

#### **Data Pre-Processing**

The following steps were done to clean the data and transform all the experiment data into a canonical form for analysis. First, the original data was retrieved for all the assignments that were known to be experiments by using the ASSISTmentsTestBed.org. Once this data

Experiment Number	Students	Problems
1	1,204	7,657
2	695	9,466
3	381	5,105
4	627	9,771
5	1,864	12,549
6	457	1,861
7	515	2,068
8	540	2,212
9	402	2,213
10	545	8,030
11	412	2,612
12	335	1,817
13	497	3,930
14	704	3,197
15	523	2,967
16	136	1,526
17	428	2,930
18	745	3,776
19	1330	4,694
20	1,917	9,268
21	429	2,538
22	426	2,065

**Table 1.** Number of students and problems done for each experiment.

was retrieved, problems which were not common to all the experiments were removed. These problems included any pretest and posttest problems that an experiment may have contained as well as scaffold problems. Nearly all the prior experiments did not have pretest or posttest problems, since the ability to create such experiments did not exist in the ASSISTments system at the time those experiments were created. Scaffold problems are sub-problems for a given original problem, which break the original problem down into a few smaller sub-problems. Scaffold problems were removed because only eleven experiments contained scaffold problems and scaffold problems are considered a type of tutoring which does not count toward the correctness of the Skill Builder assignment.

A recent improvement to the ASSISTments system was the addition of logging the condition students were randomly assigned into. Unfortunately this data did not exist at the time most of the experiments were run, therefore the condition that students were assigned into was derived from the problems they had been assigned. The condition students were randomly assigned into was derived based on distinguishing problem numbers that the student had worked on.

Several experiments involved interventions containing video feedback, where students were required to have video in order to be considered part of the experiment. These experiments had an initial question to check on whether or not the student could see the video. The answers to these questions were analyzed to determine if the student was able to see video. An additional column was added to the dataset to indicate if the student could see video where applicable.

The two variables used for measuring student performance were whether or not the student completed the assignment and the mastery speed for those students who were able to complete the assignment. The number of problems it takes a student to complete a skill builder is also called mastery speed as defined in [10]. We believe these two variables are important measures, where mastery speed measures student learning and completion rate measures student persistence. There were a few cases where a small number of students did a large number of problems, which added skew to the distribution of mastery speed. A log transform was used so that these few students did not have large impact on the means for mastery speed.

### Data Description

A description of the column names in the dataset are provided at the ASSISTments Data Dump Glossary webpage<sup>1</sup>. The dataset has 32 columns including a number of student covariates. The dataset is in the form of one row per student per assignment. Table 1 shows some basic summary statistics of the dataset, including the number of problems done for each experiment and the number of unique students in each experiment. A total of 102,252 problems were attempted by 8,297 unique students across 22 different experiments. Note that a single student could have participated in more than one experiment and not all students experienced the conditions of the experiment. This dataset can be downloaded at <https://sites.google.com/site/las2016data/data/ThisOne.xlsx>.

<sup>1</sup> <http://www.assistments-testbed.org/the-data/interpreting-your-data-v-1-0>

## **CONTRIBUTIONS, FUTURE WORK, AND CONCLUSIONS**

The contribution that this paper makes is the generation of a dataset consisting of 22 randomized controlled experiments run inside the ASSISTments online learning platform. What makes this dataset unique is that it is a combination of data from several different experiments all with a common design and common set of independent and dependent variables. We believe this dataset can be used for further analysis with a focus on finding general trends among independent variables across multiple experiments. It is currently being used in research to improve the detection of treatment effects in randomized controlled experiments [9].

## **Acknowledgements**

We thank Yang Lu for creating large portion of the infrastructure to support the ASSISTments TestBed which we used to obtain our data. We also thank Anthony Botelho for his work on adding several of the features we used to the data reporting system.

We thank multiple NSF grants (ACI-1440753, DRL-1252297, DRL-1109483, DRL-1316736, DGE-1535428 & DRL-1031398), the US Dept. of Ed (IES R305A120125 & R305C100024 and GAANN), and the ONR.

## **References**

- [1] ASSISTmentsTestBed (2015) The ASSISTmentsTestBed. A website that supports researchers funded via the Heffernan (2014) grant. Retrieved from [www.assistmentstestbed.org](http://www.assistmentstestbed.org)
- [2] Heffernan (2014) "SI2-SSE: Adding Research Accounts to the ASSISTments' Platform: Helping

Researchers Do Randomized Controlled Studies with Thousands of Students." An NSF proposal abstract

- [3] Lang, C., Heffernan, N., Ostrow, K., & Wang, Y. (2015). The Impact of Incorporating Student Confidence Items into an Intelligent Tutor: A Randomized Controlled Trial. In Santos, et al. (eds.) Proceedings of the 8th Int Conf on EDM. pp. 144-149.
- [4] Ostrow, K. S. (2015). A Multifaceted Consideration of Motivation and Learning within ASSISTments (Master's Thesis, Worcester Polytechnic Institute).
- [5] Ostrow, K. S., & Heffernan, N. T. (2015, June). The Role of Student Choice Within Adaptive Tutoring. In Artificial Intelligence in Education (pp. 752-755). Springer International Publishing.
- [6] Ostrow, K. S., Selent, D., Wang, Y., Van Inwegen, E. G., & Heffernan, N. T. (2015) Assessment of Learning Infrastructure (ALI): The Theory, Practice, and Scalability of Automated Assessment. Learning Analytics 2016.
- [7] Razzaq, L., Patvarczki, J., Almeida, S. F., Vartak, M., Feng, M., Heffernan, N. T., & Koedinger, K. R. (2009). The Assistent Builder: Supporting the life cycle of tutoring system content creation. Learning Technologies, IEEE Transactions on, 2(2), 157-166.
- [8] Selent, D., & Heffernan, N. (2015, June). When More Intelligent Tutoring in the Form of Buggy Messages Does not Help. In Artificial Intelligence in Education (pp. 768-771). Springer International Publishing.
- [9] Thanaporn, P., Selent, D., & Heffernan, N., Beck, J., Zou, J. Using a Single Model across Multiple Interventions to Improve the Detection of Treatment Effects. The 9th International Conference on Educational Data Mining on Educational Data Mining (*in submission*)
- [10] Xiong, X., Li, S., & Beck, J. E. (2013, May). Will You Get It Right Next Week: Predict Delayed Performance in Enhanced ITS Mastery Cycle. In FLAIRS Conference