

Fourth Annual Workshop on A/B Testing and Platform-Enabled Learning Research

Ritter, Steve; Heffernan, Neil; Williams, Joseph Jay; Lomas, Derek; Bicknell, Klinton; Roschelle, Jeremy; Motz, Ben; McNamara, Danielle; Baraniuk, Richard; More Authors

DOI

[10.1145/3573051.3593397](https://doi.org/10.1145/3573051.3593397)

Publication date

2023

Document Version

Final published version

Published in

L@S 2023 - Proceedings of the 10th ACM Conference on Learning @ Scale

Citation (APA)

Ritter, S., Heffernan, N., Williams, J. J., Lomas, D., Bicknell, K., Roschelle, J., Motz, B., McNamara, D., Baraniuk, R., & More Authors (2023). Fourth Annual Workshop on A/B Testing and Platform-Enabled Learning Research. In *L@S 2023 - Proceedings of the 10th ACM Conference on Learning @ Scale* (pp. 254-256). (L@S 2023 - Proceedings of the 10th ACM Conference on Learning @ Scale). Association for Computing Machinery (ACM). <https://doi.org/10.1145/3573051.3593397>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Fourth Annual Workshop On A/B Testing And Platform-Enabled Learning Research

Steve Ritter
Carnegie Learning
sritter@carnegielearning.com

Neil Heffernan
Worcester Polytechnic Institute
neilheffernaniii@gmail.com

Joseph Jay Williams
University of Toronto
williams@cs.toronto.edu

Derek Lomas
Delft University of Technology
derek@playpowerlabs.com

Klinton Bicknell
Duolingo
klinton@duolingo.com

Jeremy Roschelle
Digital Promise
jroschelle@digitalpromise.org

Ben Motz
Indiana University
bmotz@indiana.edu

Danielle McNamara
Arizona State University
dsmcnamara1@gmail.com

Richard Baraniuk
Rice University
richb@rice.edu

Debshila Basu Mallick
Rice University
debshila@rice.edu

Rene Kizilcec
Cornell University
kizilcec@cornell.edu

Ryan Baker
University of Pennsylvania
ryanshaunbaker@gmail.com

Stephen Fancsali
Carnegie Learning
sfancsali@carnegielearning.com

April Murphy
Carnegie Learning
amurphy@carnegielearning.com

ABSTRACT

Learning engineering adds tools and processes to learning platforms to support improvement research. One kind of tool is A/B testing—common in large software companies and also represented academically at conferences like the Annual Conference on Digital Experimentation (CODE). Recently, several A/B testing systems have arisen that focus on conducting research in educational environments, including UpGrade, Terracotta, and E-TRIALS. A/B testing can help improve educational platforms, yet there are challenging issues unique to conducting such work in these contexts. In response, a number of digital learning platforms have opened their systems to learning-improvement research by instructors and/or third-party researchers, with specific supports necessary for education-specific research designs. This workshop will explore challenges of A/B testing in educational contexts, how learning platforms are accelerating education research, and how empirical approaches can be used to drive powerful gains in student learning. It will also discuss opportunities for funding to conduct platform-enabled learning research.

CCS CONCEPTS

- General and reference → Cross-computing tools and techniques; Empirical studies; Experimentation;
- Applied computing → Education; Computer-assisted instruction; Interactive learning environments; E-learning; Computer-managed instruction;
- Human-centered computing → Human computer interaction (HCI); Empirical studies; HCI design and evaluation methods → Field studies

KEYWORDS

A/B Testing, Learning Engineering, Educational Technology, Digital Experimentation

ACM Reference format:

Steve Ritter, Neil Heffernan, Joseph Jay Williams, Derek Lomas, Klinton Bicknell, Jeremy Roschelle, Ben Motz, Danielle Mcnamara, Richard Baraniuk, Debshila Basu Mallick, Rene Kizilcec, Ryan Baker, Stephen Fancsali, & April Murphy. 2023. Fourth Annual Workshop on A/B Testing and Platform-Enabled Learning Research. In *Proceedings of the Tenth ACM Conference on Learning @ Scale (L@S'23), July 20–22, 2023, Copenhagen, Denmark*. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3573051.3593397>

THEME AND GOALS

The conference organizers all have deep practical experience with building learning engineering platforms for educational software. We will solicit presentations through the call for participation and, upon acceptance, organize those presentations into themes, which will form the basis of the workshop.



This work is licensed under a Creative Commons Attribution International 4.0 License.

L@S '23, July 20–22, 2023, Copenhagen, Denmark
© 2023 Copyright is held by the owner/author(s).
ACM ISBN 979-8-4007-0025-5/23/07.
<https://doi.org/10.1145/3573051.3593397>

This will be a full-day workshop. The morning session will be devoted to presentations and discussions of accepted papers. We will organize presenters into sessions addressing major themes (e.g., “communicating to the public about random-assignment experiments”), with the expectation that we will have 4-6 themes addressed during the workshop. Each presenter will have 15 minutes to present, followed by 5 minutes for questions. At the end of each theme session, a discussant will lead a panel discussion including the presenters and structured around discussion related to the general theme.

The afternoon session will consist of two sets of breakout sessions, organized by key questions. Attendees can choose which session they wish to attend. One session will focus more on A/B testing within one’s own platform; the other on how a researcher might use other platforms that are open to research from third parties. They collaborate on a Google Doc, which serves to guide a presentation to the full workshop, along with notes from the session.

In the 2022 workshop, the breakout groups addressed the following questions:

- How can we promote the idea to the public that (a) A/B testing and RCTs and (b) opening commonly used platforms to improvement research are paths to improving educational outcomes?
- Different communities use terms like “A/B Testing,” randomized field trials, RCTs, experiments, etc. Are these the same things? What language choices will create desirable engagement and avoid resistance?
- How do we communicate about the idea of incremental improvement in education? Does the public understand this? How do we contextualize A/B testing as part of this process?
- How can we ensure that the results of A/B tests are applicable to vulnerable and historically underserved populations and address equity?
- How do we talk about the “winner” of an A/B test in the educational context? How do we communicate about certainty of outcomes? How should we think about generalization to other populations and contexts?
- How do we talk to the education policy community about the value of A/B tests? How can we communicate about a balance between “doing what works” and “finding out what works better”?
- We sometimes fear that there will be objections to random assignment, which leads to A/B testing being done without public awareness. How can we be more open about continuous improvement and encourage the idea that A/B testing is a best practice that should be seen as a benefit to schools, teachers and students?
- How do we think about informed consent in educational A/B testing? In what cases should we expect to get approval from teachers, parents and the students themselves?
- How can we better include educators, administrators, students and teachers in the design and coordination of A/B tests?

BACKGROUND AND RELEVANCE TO L@S

There is no simple path that will take us immediately from the contemporary amateurism of the college to the professional design of learning environments and learning experiences. The most important step is to find a place on campus for a team of individuals who are professionals in the design of learning environments – learning engineers, if you will. [1]

Learning engineering adds tools and processes to learning platforms to support improvement research [2]. One kind of tool is A/B testing [3], which is common in large software companies and also represented academically at conferences like the Annual Conference on Digital Experimentation (CODE). A number of A/B testing systems focused on educational applications have arisen recently, including UpGrade [4] and E-TRIALS [5]. A/B testing can be part of the puzzle of how to improve educational platforms, and yet challenging issues in education go beyond the generic paradigm. For example, the importance of teachers and instructors to learning means that students are not only connecting with software as individuals, but also as part of a shared classroom experience. Further, learning in topics like mathematics can be highly dependent on prior learning, and thus A or B may not be better overall, but only in interaction with prior knowledge [6]. In response, a set of learning platforms is opening their systems to improvement research by instructors and/or third-party researchers, with specific supports necessary for education-specific research designs. This workshop will explore how A/B testing in educational contexts is different, how learning platforms are opening up new possibilities, and how these empirical approaches can be used to drive powerful gains in student learning. It will also discuss forthcoming opportunities for funding to conduct platform-enabled learning research.

Issues to be addressed are expected to include:

- The role of A/B testing systems in complying with SEER principles (<https://ies.ed.gov/seer/>), which set a high bar for the goals of empirical studies of educational improvement
- Awareness of opportunities to use existing learning platforms to conduct research (<http://seernet.org>)
- Managing unit of assignment issues, as arise when students are in classrooms with a shared teacher
- Practical considerations related to experimenting in school settings, MOOCs, & other contexts
- Ethical, data security, and privacy issues
- Relating experimental results to learning-science principles
- Understanding use cases (core, supplemental, in-school, out-of-school, etc.)
- Accounting for interactions between the intended contrast (A vs. B) and learners’ prior knowledge, aptitudes, background, or other important variables
- A/B testing within adaptive software
- Adaptive experimentation

- Attrition and dropout
- Stopping criteria
- User experience issues
- Educator involvement and public perceptions of experimentation
- Balancing practical improvements with open and generalizable science

The 2020, 2021, and 2022 workshops on this topic were very successful, with some of the highest registrations of any workshops at the conference. We welcome participation from researchers and practitioners who have either practical or theoretical experience related to running A/B tests and/or randomized trials as well as platform-enabled learning research. This may include researchers with backgrounds in learning science, computer science, economics and/or statistics.

EXPECTED OUTCOMES AND CONTRIBUTIONS

We will publish papers and continue to develop and deploy systems in this area. Following the earlier workshops, we created a Slack channel for continuing discussions on this issue (<https://edu-ab-testing.slack.com/>), and plan to continue discussions on this forum. We expect to conduct this workshop yearly, as it becomes part of the basis for a community of researchers conducting A/B tests at scale.

REFERENCES

- [1] Herbert A. Simon. 1967. The job of a college president. *Educational Record*, 48, 68-78.
- [2] Melina R. Uncapher (2018): From the science of learning (and development) to learning engineering, *Applied Developmental Science*, <https://doi.org/10.1080/10888691.2017.1421437>
- [3] Kohavi, R., Deng, A., Frasca, B., Walker, T., Xu, Y., & Pohlmann, N. (2013, August). Online controlled experiments at large scale. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1168-1176).
- [4] Ritter, S., Murphy, A., Fancsali, S. Lomas, D., Fitkariwala, V. and Patel, N.. (2020). UpGrade: An open source tool to support A/B testing in educational software. *L@S Workshop on A/B Testing at Scale*.
- [5] Ostrow, K.S., Heffernan, N.T., & Williams, J.J. (2017). Tomorrow's EdTech Today: Establishing a Learning Platform as a Collaborative Research Tool for Sound Science. *Teachers College Record*, Volume 119 Number 3, 2017, 1-36.
- [6] Fyfe, E. R. (2016). Providing feedback on computer-based homework in middle-school classrooms. *Computers in Human Behavior*, 63, 568-574. <https://doi.org/10.1016/j.chb.2016.05.082>