

Beyond Failure: The 2nd LAK Failathon Poster

Doug Clow¹, Rebecca Ferguson¹, Kirsty Kitto², Yong-Sang Cho³

¹The Open University
Walton Hall, Milton Keynes
MK7 6AA, UK
Firstname.Surname@open.ac.uk

²Queensland University of
Technology, School of Mathematical
Sciences, Level 7, P-Block, 2 George
Street, Brisbane, 4001, Australia
kirsty.kitto@qut.edu.au

³Korea Education & Research
Information Service
64 Dongnau-Ro, Dong-Gu, Daegu,
41061 Korea
zzosang@gmail.com

ABSTRACT

This poster will be a chance for a wider LAK audience to engage with the 2nd LAK Failathon workshop. Both of these will build on the successful Failathon event in 2016 and extend beyond discussing individual experiences of failure to exploring how the field can improve, particularly regarding the creation and use of evidence.

Failure in research is an increasingly hot topic, with high-profile crises of confidence in the published research literature in medicine and psychology. Among the major factors in this research crisis are the many incentives to report and publish only positive findings. These incentives prevent the field in general from learning from negative findings, and almost entirely preclude the publication of mistakes and errors. Thus providing an alternative forum for practitioners and researchers to learn from each other's failures can be very productive. The first LAK Failathon, held in 2016, provided just such an opportunity for researchers and practitioners to share their failures and negative findings in a lower-stakes environment, to help participants learn from each other's mistakes. It was very successful, and there was strong support for running it as an annual event. The 2nd LAK Failathon workshop will build on that success, with twin objectives to provide an environment for individuals to learn from each other's failures, and also to co-develop plans for how we as a field can better build and deploy our evidence base.

This poster is an opportunity for wider feedback on the plans developed in the workshop, with interactive use of sticky coloured dots to illustrate. This broadens the participant base in this important work, which should improve the quality of the plans and the commitment of the community to delivering them.

Categories and Subject Descriptors

K.3.1 [Computers and Education]: Computer Uses in Education

General Terms

Management, Human Factors.

Keywords

Learning analytics, analytics, evidence, learning from failure.

SAMPLE: Permission to make digital or hard copies of all or part 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. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '10, Month 1–2, 2010, City, State, Country.

Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

DOI: <http://dx.doi.org/10.1145/12345.67890>

1 WORKSHOP BACKGROUND

1.1 Failure in research

Problems with the published research literature are currently receiving large amounts of attention, particularly in applied fields.

In health, the optimism that surrounded the evidence-based medicine movement is beginning to falter, partly as the idea is diverted from its original goals [1], but more fundamentally, as issues with the underlying research come to light. Not only is most published research false [2], but most of the true research that is published is not useful in clinical practice [3].

In psychology, the 'replication crisis' continues and intensifies. A prominent effort to replicate a series of 100 classic psychological results [4] achieved very partial success: "A large portion of replications produced weaker evidence for the original findings", with only 36–47% of replications succeeding, depending on the measure chosen. It has also proved highly controversial, with many blog and social media posts, using language that is sometimes intemperate. One recent high-profile example of a failed replication is 'power poses'. The original claim was that "a person can, by assuming two simple 1-min poses, embody power and instantly become more powerful" [5]. One of the original authors has had significant success as a public speaker on the topic, with a TED talk receiving over 36m views [6], but after failed replications, one of the authors has very creditably concluded that they "do not think the effect is real" [7].

A wide range of complex and hard-to-overcome factors lies behind these problems in establishing a strong evidence base for practice. Many of these concern the use of statistics, including the use of 'researcher degrees of freedom' to achieve significance [8] – importantly, this is not limited to situations where researchers conduct multiple unreported comparisons, but also where 'researchers can perform a reasonable analysis given their assumptions and their data, but had the data turned out differently, they could have done other analyses that were just as reasonable' [9]. Fundamentally, any research carried out with low pre-study odds is prone to false positives [2]. Incentives on researchers to publish significant findings play a strong part, and may encourage publication of low-quality research even if replications were commonplace and there were significant negative consequences to publishing studies that were later repudiated [10].

The 'file drawer' effect, whereby uninteresting or negative findings are not reported, is a major concern. In clinical research, the ambitious AllTrials¹ project seeks to ensure "All trials registered, all results reported" to reduce this problem.

¹ <http://www.alltrials.net/>

1.2 Evidence in learning analytics

There is no reason to believe that learning analytics is immune to these problems. One attempt to explore this issue is the Learning Analytics Community Exchange (LACE) project's Evidence Hub², which maps research evidence against four propositions about learning analytics. The great majority of evidence classified was positive, with only 14% negative [11], which suggests that there is a significant publication bias in the field. Further, very little of the published research could be classified at the 'higher' levels of the evidence hierarchy (i.e. systematic reviews, randomised controlled trials) [11]. These are the base levels at which the problems in health and psychology can be detected, so their dearth in the evidence base for learning analytics may mean that the problems in our field are even more profound.

1.3 Why a workshop at LAK

The first LAK Failathon was a success, giving an opportunity for practitioners and researchers to talk about – and learn from – their failures in a way that is difficult to provide in any other context. This second LAK Failathon will build on that success and provide a similar space in the first half of the workshop.

The critiques in health and psychology propose a wide range of possible solutions (e.g. [12]), some of which may well be useful in the field of learning analytics. So the second part of the workshop will explore, collectively, how we can improve the creation and use of evidence in our field.

2. WORKSHOP OBJECTIVES AND INTENDED OUTCOMES

This workshop has two chief objectives: firstly, to provide an effective space for sharing experiences of failures, and secondly, to work collaboratively to produce prioritised action plans for the field of learning analytics to improve.

2.1 Sharing experience of failure

The first part of the workshop aims to allow practitioners and researchers to learn from each other's mistakes. There are strong pressures on people to publicise success and minimise failures, which limit the willingness of people to admit their mistakes and discuss them. Closed forums are routinely used in education, in part to allow learners to have a safe space to make mistakes, from which they can learn. So, as with last year, this part of the workshop will be held under the Chatham House Rule:

“When a meeting, or part thereof, is held under the Chatham House Rule, participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.”³

The participants can discuss, write and change their plans based on what they learn from this part of the workshop, but may not identify who said it or which organisation or activity it related to.

2.2 Producing action plans for improvement

It is helpful to learn as individuals from each other's mistakes, but it is also helpful to learn and improve collectively. This year's Failathon is focused particularly on evidence, and this part of the workshop aims to explore what can be done to improve the creation and use of evidence in the field of learning analytics.

The chief outcome from the workshop will be a series of action plans collectively developed by the participants, consisting of prioritised lists of suggested actions that could be taken by:

- SoLAR, the Society for Learning Analytics Research
- Future LAK conference organisers and committees
- Universities and other research organisations
- Companies, developers, and others with interests in learning analytics

3. Poster session

Following the workshop, we will take the plans developed by the participants to the LAK poster session, to solicit feedback from a broader audience, via coloured dots (green / yellow / red to indicate support / caveat / oppose proposals) and sticky notes (green for new ideas, yellow for comments). This will engage the community more broadly than the workshop participants, which will raise the profile of these issues, and give the plans as finally developed greater legitimacy and, one may hope, traction.

4. REFERENCES

- [1] Greenhalgh, T., Howick, J., & Maskrey, N. (2014). Evidence based medicine: a movement in crisis? *BMJ* 2014;348:g3725
- [2] Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med*, 2(8), e124.
- [3] Ioannidis, J.P.A. 2016. Why Most Clinical Research Is Not Useful. *PLoS Med* 13(6): e1002049.
- [4] Open Science Collaboration. (2015) Estimating the reproducibility of psychological science. *Science*, 28 Aug 2015: 349(6251).
- [5] Carney, D. R., Cuddy, A. J., & Yap, A. J. (2010). Power posing brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological Science*, 21(10), 1363-1368.
- [6] Cuddy, A. 2012. Your body language shapes who you are. *TED talk*, https://www.ted.com/talks/amy_cuddy_your_body_language_shapes_who_you_are
- [7] Carney, D. 2016. *My position on "Power Poses"*. Blog post: http://faculty.haas.berkeley.edu/dana_carney/pdf/My%20position%20on%20power%20poses.pdf
- [8] Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 0956797611417632.
- [9] Gelman, A. & Loken, E. (2013) *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*. Blog post: http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf
- [10] Smaldino, P.E. & McElreath, R. 2016 The natural selection of bad science. *R. Soc. open sci.* 3:160384.
- [11] Ferguson, R & Clow, D. (2015) *Evidence Hub Second Review D2.8*. <http://www.laceproject.eu/deliverables/d2-8-evidence-hub-second-review/>
- [12] Ioannidis JPA (2014) *How to Make More Published Research True*. *PLoS Med* 11(10): e1001747. doi:10.1371/journal.pmed.1001747

² <http://evidence.laceproject.eu/>

³ <https://www.chathamhouse.org/about/chatham-house-rule>

Columns on Last Page Should Be Made As Close As Possible to Equal Length