A Focus on Methodology in Learning Analytics: Building a Structurally Sound Bridge Discipline

Yoav Bergner
New York University

Charles Lang
Teachers College
Columbia University

Geraldine Gray
Institute of Technology, Blanchardstown

Abstract

The following paper gives an overview of the inaugural Methodologies in Learning Analytics Workshop at the International Learning Analytics & Knowledge Conference 2017 in Vancouver, Canada. The event discussed many topics but two key themes emerged, that of middle space, the space between learning and analytics in which methodologies reside and the importance of multivocality, the challenge of finding shared analytic objectives and learning to not talk past one another. The following summarizes these themes and their importance for generating robust methodological arguments within Learning Analytics.

Introduction

In the third year of the International Learning Analytics & Knowledge Conference (LAK2013), the conference organizers sketched out a theme of dialectics in learning analytics [15]. One of these dialectics was middle space—as in, a space in between learning and analytics. Another was productive multivocality—that is, finding shared analytic objectives and learning to not talk past each other. In preparing our workshop proposal for LAK2017, we found ourselves wanting to push a bit harder on the conceptual infrastructure necessary for sustaining this middle space and for effective communication, or what it means to build a structurally sound bridge discipline.

From its beginnings, learning analytics emphasized the importance of bridging computer sciences and social sciences [13]. Learning analytics has also been described as helping to bridge education, psychology, and neuroscience [10]. Interestingly, in Boyack, Klavans, and Börner's [3] scientometric analysis of the "backbone of science," a network analysis based on journal inter-citations found both education and computer science lie toward the outside of the network. Education and computer science are fairly insular by the authors' analysis, while artificial intelligence and psychology are more central. The shortest path connecting education to computer science passes through psychology and statistics. Thus, one might speculate that if Boyack et al. were to redo their analysis in the future, learning analytics and educational data mining journals would cluster somewhere on this path.

Mindful of Suthers and Verbert's second dialectic, we see productive multivocality as intimately connected to methodology and the uses of statistics as principled argument (cf. [1]) to support claims about learning and educational improvement. Methodologists are the quality assurance engineers in the bridge-building enterprise of learning analytics, stress testing the roadways, trusses, and wire ropes that connect educational technologists, psychologists, data scientists, learning scientists, substantive experts in various educational domains, and measurement specialists. For all of the strength that comes from such diversity of expertise, there are also challenges when it comes to establishing norms for methodological work. Learning analytics has embraced an eclectic approach to methodology but may lack its own coherent epistemology [5]. In an *amicus brief* of a paper to his own scholarly community, Peter Kennedy [7] enumerated the ten commandments of applied econometrics, that is, the unwritten rules that guide methodological choices. This short paper does not attempt to establish the rules of learning analytics methodology, but only to make the case that some set of rules is desirable.

Papers as Arguments

To elucidate the role of a methodological focus in learning analytics, consider the general framework for reasoning or argument due to Toulmin [16]. A representation of the Toulmin model is shown in Figure 1. The end of an argument is a claim, which is supported along a central path by data/observations. However, the observations alone do not suffice to model non-trivial arguments, which usually involve some warrant for justifying the claim from the observations. The warrant itself is backed by some other observations. Finally, alternative explanations and rebuttal evidence serve to qualify a claim or, possibly, strengthen it if alternative explanations appear to be weak. The model is simple but remarkably general.

An example of Toulmin's model in the context of a simple school assessment might go as follows: "Donald is very weak in mathematics" (claim). "He got almost all of the questions wrong on the diagnostic test" (data). "These questions were a good gauge of grade 6 math ability" (warrant): "they were drawn from a pool that our school developed with expert help and refined over several years. Math teachers say the scores on the placement test indeed predict which students struggle without extra help and which students succeed" (backing). Or, "This math test did not accurately gauge Donald's math ability" (alternative explanation): "Donald suffers from health problems that kept him awake throughout the night, and he was too exhausted to perform at his actual level" (rebuttal evidence).
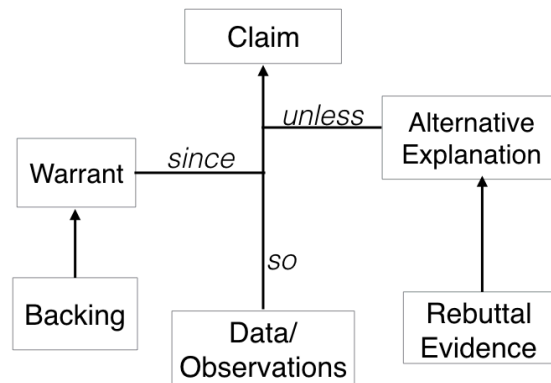
Figure 1: Toulmin model of a reasoning argument.

Of course, a paper is usually composed of a chain of arguments rather than a single one. A simplified representation of an archetypal learning analytics paper might involve the chain illustrated in Figure 2. It is not implied that the paper is written in this order, but rather that the logical path from data/observations to a claim (about learners, technologies, pedagogies, etc.) will typically involve warrants justifying data selection and preparation (possibly multiple such steps), model selection and implementation, and evaluation. For each of these links in the chain of argument, alternative explanations might exist that potentially undermine the claim. For example, was the data selection or transformation justifiably warranted given the claim and the data? Paper authors who aim to build a strong argument are likely to engage with these alternative explanations. Methodologists, in particular though, make it their business to understand the underside in this diagram. This is hardly to suggest that methodologists are a finger-wagging bunch who delight in niggling their colleagues about violations of model assumptions. Alternative explanations and warrants go hand in hand in building an argument. The exhaustion of alternative explanations can strengthen the warrant.
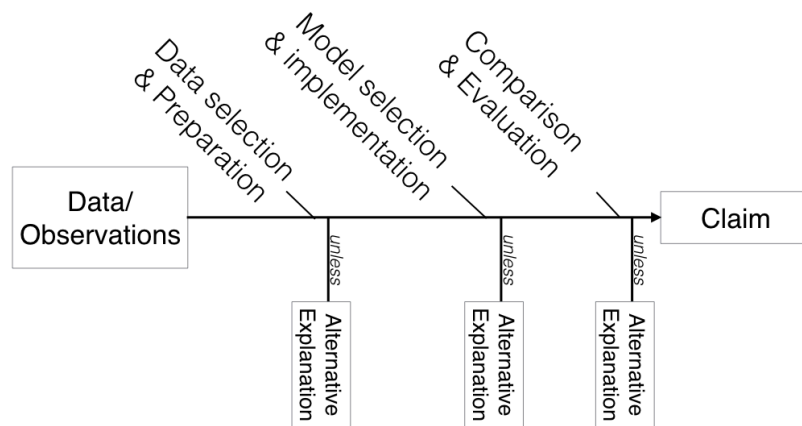

Figure 2: Sample chain of arguments in a learning analytics paper

Chaining weak links undermines the structural integrity of an argument, in learning analytics as anywhere. However, learning analytics may be particularly susceptible to this weakness given the breadth of techniques practitioners use. It is challenging for readers and reviewers to be fluent with all of them. We will cherry pick an example from one of the pioneering papers on modeling student engagement in massive open online courses by Kizilcec, Piech, and Schneider (KPS; [8]). We single this paper out not because it represents a particularly egregious example, but only because it is one of the most cited references[1] in the learning analytics literature supporting the discrete characterization of MOOC learners by disengagement patterns.

KPS arrive at the plausible claim that MOOC learners can be categorized as completing, auditing, disengaging, or sampling. On their way to this claim, the following three steps are involved. First, in each assessment period, learners are labeled as "on track" (did the assessment on time), "behind" (turned in the assessment late), "auditing" (didn't do the assessment but engaged by watching a video or doing a quiz), or "out" (didn't participate in the course at all), leading to a vector of engagement observations, for example, [T, T, T, T, T, B, A, A, A]. In the second step, the similarity between engagement vectors for two students is computed as follows: assign numerical values to each label (on track = 3, behind = 2, auditing = 1, out = 0), and compute the L1 norm of the list of numbers. Finally, in step 3, k-means clustering is used (repeated 100 times from random start points). What is the problem with this sequence of steps?

The first step in the above sequence involves (several instances of) dichotomizing a continuous variable, a practice generally frowned upon for increasing the likelihood of type I and type II errors [11]. Does it matter how late is late? Or whether a student watched one video or ten? Perhaps not, but the authors do not make this case. In the second step, a categorical label is transformed into an interval scale for the purpose of computing distances. Is this justified? (Assumes the "difference" between two learners is the same if (a) one of them watched 10 videos and the other none or (b) if one of them completed an assignment on time and the other completed it late.) Lastly, the use of k-means is suspect with a non-Euclidean metric as is used in the KPS analysis; a k-medians modification is recommended for L1 norms [14]. In summary, the KPS analysis involves a chaining of three analytical steps, each of which is potentially suspect. Does this mean the ultimate claim is wrong? Of course not. However, we note that not only has this paper been cited for its claims about learners, but its methods have also been used in replication studies (e.g., [6]). Repeated use in itself becomes evidence for validity and tacitly vindicates the lack of consideration of alternative explanations.

Operationalization and Sensitivity Analyses

What KPS perhaps lacked most was an analysis of the sensitivity of results to operationalization of their engagement variables and distance metrics. We are reminded of Kennedy's [7] tenth commandment of applied econometrics: "*Thou shalt confess in the presence of sensitivity* (Corollary: Thou shalt anticipate criticism)" (p. 583). In fact, as the field of

---

[1] At the time of writing, this paper was cited over 500 times according to Google Scholar.

learning analytics has matured, a number of more recent papers have emphasized the sensitivity of quantitative analyses to data collection and variable operationalization choices, for example in the cases of selection bias [4], time-on-task analyses [9], studies of discussion forum usage [2], and evaluation of student models [12]. Our aim is not to ring an alarm bell. At the risk of stating the obvious, we emphasize only that the methods we use in learning analytics are subject to random and systematic error. If we do not make explicit efforts to quantify uncertainty of both kinds, we chain together weak links.

Conclusion

In conclusion, the Inaugural Workshop in Methodologies in Learning Analytics raised more questions than it answered, but we are confident that this is a positive sign. It demonstrates the thirst of the community to engage critically in methodological conversations and address the challenges of building methodological bridges within the discipline. The form that this endeavor will ultimately take within the community will largely depend on the ideas discussed, middlespace and multivocality. The challenge is to define the objectives of the field, align those objectives with methodologies, and communicate those arguments across the many fields involved in learning analytics and beyond.

Acknowledgement

References

[1]   Abelson, R.P. 1995. *Statistics as principled argument*. L. Erlbaum Associates.

[2]   Bergner, Y. et al. 2015. Methodological Challenges in the Analysis of MOOC Data for Exploring the Relationship between Discussion Forum Views and Learning Outcomes. *Proceedings of 8th International Conference on Educational Data Mining*. (2015).

[3]   Boyack, K.W. et al. 2005. Mapping the backbone of science. *Scientometrics*. 64, 3 (Aug. 2005), 351–374.

[4]   Brooks, C. et al. 2015. Reducing selection bias in quasi-experimental educational studies. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (2015), 295–299.

[5]   Clow, D. 2013. An overview of learning analytics. *Teaching in Higher Education*. 18, 6 (2013), 683–695.

[6]   Ferguson, R. and Clow, D. 2015. Examining engagement. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*. (2015), 51–58.

[7]   Kennedy, P.E. 2002. Sinning in the basement: what are the rules? The ten commandments of applied econometrics. *Journal of Economic Surveys*. 16, 4 (2002),

569–589.

[8]   Kizilcec, R. et al. 2013. Deconstructing Disengagement : Analyzing Learner Subpopulations in Massive Open Online Courses. *Proceedings of the Third International Conference on Learning Analytics and Knowledge Discovery*. (2013).

[9]   Kovanović, V. et al. 2015. Penetrating the black box of time-on-task estimation. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15* (New York, New York, USA, Mar. 2015), 184–193.

[10]   Lodge, J.M. and Corrin, L. 2017. What data and analytics can and do say about effective learning. *npj Science of Learning*. 2, 5 (2017), 1–2.

[11]   MacCallum, R.C. et al. 2002. On the practice of dichotomization of quantitative variables. *Psychological methods*. 7, 1 (2002), 19–40.

[12]   Pelánek, R. et al. 2016. Impact of data collection on interpretation and evaluation of student models. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (2016), 40–47.

[13]   Siemens, G. 2012. Learning Analytics : Envisioning a Research Discipline and a Domain of Practice. *2nd International Conference on Learning Analytics & Knowledge*. May (2012), 4–8.

[14]   Steinley, D. 2006. K-means clustering: a half-century synthesis. *The British journal of mathematical and statistical psychology*. 59, Pt 1 (May 2006), 1–34.

[15]   Suthers, D. and Verbert, K. 2013. Learning analytics as a middle space. *International Conference on Learning Analytics 2013*. (2013), 2–5.

[16]   Toulmin, S.E. 2003. *The Uses of Argument (Updated edition 2003)*. Cambridge University Press.