Mathematical Values in Data Science: Three Conjectures to Explain Mathwashing

Patrick Allo

With the development of a critical research agenda on contemporary data practices we gradually build the tools that are needed to overcome the uncertainty, lack of clarity, and impact of misleading narratives concerning the epistemology of data science. Without such a reflection, we cannot understand the kind of knowledge data analysis produces. More importantly, we then also lack the ability to evaluate specific knowledge-claims as well as more general affirmations of the epistemic superiority (smarter, more objective, ...) of the knowledge, decisions, or insights that data analysis produces. This is why it is important to recognise that data is never just data (e.g. Gitelman 2013, Kitchin 2014), or that the development of algorithms (as any advanced scientific or engineering practice) cannot fully be understood in terms of a well-defined internal logic.

The starting point of this contribution is that we should start asking similar questions about mathematics: We need to understand how mathematics contributes to scientific respectability and authority of data science. To do so, we cannot limit our attention to mathematics as a body of mathematical truths or mathematical techniques. Instead, we should focus on mathematical thought and beliefs about the nature of mathematical thought. I propose to develop this critical inquiry through a dedicated consideration of how *mathematical values* shape data science.

The study of mathematical values such as objectivity and certainty, but also the beauty of proof or the independent verifiability of calculations has a long history within mathematics education (Ernest 2016, Bishop 1991). But the study of mathematical values can also help us to understand its influence on science and society. A closer look at these values can help us understand why we consider some types of knowledge to be more deserving than others. Perhaps, it can even help us to diagnose why some types of knowledge or some classes of knowledge claims can more easily withstand criticism than others. This ability is exactly what we require of a critical epistemology of data science.

By adopting a value-oriented perspective, I do not only underscore that mathematics might not be value-free (which is in itself sufficient to warrant a closer examination of how mathematics contributes to the credibility of data science), but I also introduce a richer vocabulary to characterise the workings of mathe-

matics or to explain the emergence of hard to criticise knowledge-claims. In that sense, a value-oriented perspective can help us to uncover the epistemic mechanisms that make mathwashing—the unwarranted inference from the reliance on mathematical techniques to the neutrality of outcomes—possible.

In his study of the emergence of statistical reason, Desrosières (2010) argues that scientific practices such as statistics cannot be reduced to their supposed internal logic, but are also shaped by properties they inherit from other practices. If so, it is not unreasonable to suppose that, through its reliance on mathematical techniques, data science will inherit or be shaped by certain mathematical values. I propose three complementary precisifications of this idea. For now, let us just think of these as conjectures.

- If we think of a mathematical value, such as objectivity, as something worth pursuing, then the claim that a mathematical value can be inherited by data science implies that within data science we will also value objective knowledge more highly than subjective knowledge.
- If we think of a mathematical value as a property of mathematical knowledge, such as certainty, then the claim that a mathematical value can be inherited by data science implies that we will tend to attribute the same property (in this example, certainty) to the epistemic achievements of data science.
- If we think of a mathematical value as a criterion that distinguishes good from bad mathematics, such as clarity or precision, then the claim that a mathematical value can be inherited by data science implies that the same criterion could or ought to be used to evaluate data science.

These conjectured mechanisms of value-inheritance point to potential reasons for trusting data science, for seeing the development of data science as an epistemologically positive evolution, or for the difficulties one might encounter when challenging certain data-driven knowledge-claims. By bringing these reasons into focus, we can evaluate them and ask when mathematical values serve a critical purpose, and when they do the exact opposite and make critique impossible.

References

- Bishop, Alan J. 1991. *Mathematical Enculturation : A Cultural Perspective on Mathematics Education.* Kluwer Academic Publishers.
- Desrosières, Alain. 2010. *La Politique Des Grands Nombres Histoire de La Raison Statistique.* 2nd ed. La Découverte.
- Ernest, Paul. 2016. "Mathematics and Values." In *Mathematical Cultures. The London Meetings 2012-2014.*, edited by Brendan Larvor, 189–214. Cham: Springer International Publishing.
- Gitelman, Lisa, ed. 2018. *"Raw Data" Is an Oxymoron.* Cambridge, Ma: MIT Press.

- Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences.* Sage.