

editorial: epistemologies of data

On Education

This issue of On_Education addresses data as the foundational means of research practice, ethical and social implications of data, as well as the governance of data by highlighting the recent past of data (and its bias), methodological implications of data, ontologies of data, and the political relevance of data. The seven articles offer thought-provoking perspectives to deepen our understanding of the complex relationship between data and educational research. With these perspectives, the editors hope to contribute to and stimulate ongoing discussions about the meaning and relevance of data use in and for education.

Keywords: data, editorial, epistemology

Data analysis has always been part of scientific research. For example, while the PISA assessments have significantly informed the educational discourse and the media since the turn of the twenty-first century, the collection and international comparison of quantitative data on a large scale has a more extended history: the International Association for the Evaluation of Educational Achievements (IAE) was founded in 1958 due to a lack of comparative data on student performance (Aljets, 2015; Kyriakides & Charalambous, 2014). But something seems different in the digitally driven data-hype of today: with the arrival of ‘Big Data’, ‘Artificial Intelligence’ and corresponding software, data has become even more ubiquitous. Some observers even point out that data has become so pervasive that it constructs our (digitally mediated) reality in previously unseen ways (Couldry & Hepp, 2016; Hepp, 2019). Research objects are increasingly based on digital constructs, and so is research practice.

This issue of On_Education deals with the past, present, and future of data use. It aims to shed light on datafication in and of educational research, focusing on processes, materialities and human-non-human relationships. We address epistemological implications of datafication for educational research by focusing on how data affects studies in terms of decisions, how these decisions in turn shape research practices and subsequent scholarly knowledge, and what the policy effects of that knowledge might be. The issue considers data as the foundational means of research practice, the ethical and social implications of data, as well as the governance of data. The seven articles included in the issue address different aspects of epistemologies of data that we have organised into four thematic sections:

The recent past of data (and its bias). While ‘datafication debates’ almost exclusively focus on ‘the digital’, the gathering, exchange, and analysis of data came before the digital turn: Paper-based ‘analogue’ data collection was an important basis for population statistics in the nineteenth century, and punched-card data processing became relevant to scientific inquiry in the twentieth century (Agar, 2006; Oertzen, 2017). But (how) has the use of data changed due to the introduction of

digital computers?

The contribution of Joakim Landahl presents challenges posed to comparative educational research in the early days of computer use in the 1960s, when their numbers were still small. Using computers in standardised international studies led to the problem that data still had to be distributed worldwide in a physical form, running the risk of data loss or destruction on the way. While such cases of data imperfection may be addressed with the increasing number of computers and with the internet infrastructure allowing for decentralised computer use, Landahl invites us to reflect on the precarious status of knowledge even today, as many decisions still precede the results in standardised reports that contain potential biases due to unintended external conditions.

Methodological implications of data. Critical studies of data use suggest that ‘data science’ is not a set of ‘neutral’ tools; instead, it is imbued with specific descriptive taxonomies and normative assumptions (Sætra, 2018; Desai et al., 2022). Data are often assigned values such as ‘reliability’ and ‘transparency’, but their collection can be considered as an active mediator affecting the epistemic process itself (van Es et al., 2001). At the same time, digital data methodologies bring with them the promise of rapid ‘delivery’ of ‘evidence’ as well as that of rendering the invisible visible. What are the consequences of such promises for critical and reflective educational research?

Rolf Strietholt and Stefan Johansson explore critical design factors that influence data quality, focusing on international large-scale educational assessments. The authors delve into statistical challenges related to sampling strategies, the relevant measurement issues, and the various causal conclusions derived from such assessments. Their study aims to foster a deeper understanding of the implications of data quality and design choices in comparative studies.

Thomas Hillman, Felicitas Macgilchrist, and Svea Kiesewetter pursue a similar aim of interrogating design decisions from a self-reflexive perspective. They focus on InfraReveal, a data visualisation tool developed by the authors to illustrate global infrastructures underlying educational data circulation. By reflecting on how this tool (re)produces false certainties but also geographies of power and digitalities, they raise key epistemological and methodological questions about the role of critical research in knowledge production. They propose strategies of ‘grappling with’ rather than ‘overcoming’ or ‘negating’ the inherent tensions of data visualisation.

Irina Zakharova and Annekatrin Bock focus on the in/visibilising role of data practices to reflect on the methodological consequences of researching datafication in education. Based on an overview of literature on digital technologies and datafication, they distinguish four modes of in/visibilities produced by data practices and discuss various methodological strategies employed by critical scholars in dealing with the difficult task of approaching in/visibility. They argue that software studies, critical data studies, and studies on the materiality of data have opened promising methodological avenues.

Ontologies of data. Data and its technical processing are often perceived as being able to provide (more) objective information and are less prone to subjective interpretation. However, both data collection and the development of automated processing mechanisms (i.e., software, algorithms) are also products of human beings, and hence also contain errors through design and by designers (Badke-Schaub et al., 2012). At the same time, the human component is also discussed in terms of its potential to minimise risks and dangers, given that machines likewise contain biases (Bowie & Jeffcott, 2016).

In that regard, Vassilis Galanos raises questions about the ontological nature of data by presenting and developing a philosophical concept of many-sidedness into seemingly contradictory arguments to discuss data in education. The argumentation places data, their construction, and in particular their meaning (and meaninglessness) in a new perspective relevant for future critical debates on what data are and what their use aims at.

The political relevance of data. As the need for policymaking that is both quickly produced and evidence-based—therefore relying on data—grows, debates on the politics of data are gaining new impetus. The use of technology to gather, analyse, and manage data is deeply intertwined with politics (Hepp et al., 2022), and given the need for data storage and sharing, it also has a significant environmental impact (Lucivero, 2020; Monserrate, 2022). How is the increasing requirement for ‘evidence-based’ research changing how research ideas come into being in the first place, let alone how research plans and decisions are made?

Michael Brock and Franciska Mahl examine the interplay between scientific and political-administrative perspectives in the context of municipal educational monitoring. Their contribution highlights the growing influence of data-driven approaches in educational policymaking. By analysing cases using contextual analysis, the authors discuss how these monitoring practices are not only shaped by scientific rigour but also by political considerations. They further explore how the interplay between evidence, interests, and power shapes educational governance.

Finally, Sigrid Hartong scrutinises large-scale data infrastructures by focusing on the spread of *Research Data Management (RDM)* programs, policies, and practices. While these programmes are meant to ‘open up’ science, Hartong argues that RDM inadvertently ‘close it down’. RDM raise epistemological and political questions about narrowing down what can/should be known and what can/should be funded; and they subject researchers to increasing dataveillance, shareveillance, and monitoring practices. She highlights the dangers and ethical concerns raised by the predominantly ‘solutionist-positivist’ perspectives embedded in current discourses of RDM and sketches alternative strategies to produce ‘reflective cracks’ in infrastructure systems that could minimise these risks.

Altogether, this issue of *On_Education* aims to contribute to the ongoing, critical debates about the status and significance of knowledge production in educational research and beyond. By offering a variety of perspectives on ‘epistemologies of data’, we want to invite our readers to enter this discussion and encourage them to share their perspectives on the topic.

References

Agar, J. (2006). “What difference did computers make?” *Social Studies of Science*, 36(6), 869–907.
<https://www.jstor.org/stable/25474487>

Aljets, E. (2015). *Der Aufstieg der Empirischen Bildungsforschung*. Springer VS.
<https://doi.org/10.1007/978-3-658-08115-7>

Badke-Schaub, P., Hofinger, G., & Lauche, K. (Eds.) (2012). *Human Factors: Psychologie sicheren Handelns in Risikobranchen* (2nd edition). Springer.

Bowie, P., & Jeffcott, S. (2016). Human factors and ergonomics for primary care. *Education for*

Primary Care, 27(2), 86–93.

<https://doi.org/10.1080/14739879.2016.1152658>

Couldry, N., & Hepp, A. (2016). *The mediated construction of reality*. Polity.

Desai, J., Watson, D., Wang, V., Taddeo, M., & Floridi, L. (2022). The epistemological foundations of data science: A critical review. *Synthese*, 200(6), 469.

<https://doi.org/10.1007/s11229-022-03933-2>.

van Es, K., Schäfer, M. T., & Wieringa, M. (2001). Tool criticism and the computational turn. A ‘methodological moment’ in media and communication studies. *M&K Medien & Kommunikationswissenschaft* 69, 46–64.

<https://doi.org/DOI:10.5771/1615-634X-2021-1-46>.

Hepp, A. (2019). *Deep mediatization*. Routledge.

Hepp, A., Jarke, J., & Kramp, L. (2022). New perspectives in critical data studies: The ambivalences of data power—An introduction. In A. Hepp, J. Jarke, & L. Kramp (Eds.), *New perspectives in critical data studies. Transforming communications – Studies in cross-media research* (pp. 1–23). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-96180-0_1

Kyriakides, L & Charalambous, Ch. Y. (2014). Educational effectiveness research and international comparative studies: Looking back and looking forward. In Strietholt, R., Bos, W., Gustafsson, J.-E., & Rosèn, M. (Eds.), *Educational Policy Evaluation through International Comparative Assessment* (pp. 33-49). Waxmann.

Lucivero, F. (2020). Big data, big waste? A reflection on the environmental sustainability of Big Data initiatives. *Science and Engineering Ethics*, 26(2), 1009–1030.

<https://doi.org/10.1007/s11948-019-00171-7>

Monserrate, S.G. (2022). The cloud is material: On the environmental impacts of computation and data storage. *MIT Case Studies in Social and Ethical Responsibilities of Computing*.

<https://doi.org/10.21428/2c646de5.031d4553>

Sætra, H. S. (2018). Science as a vocation in the era of Big Data: The philosophy of science behind Big Data and humanity’s continued part in science. *Integrative Psychological and Behavioral Science*, 52(4), 508–22. <https://doi.org/10.1007/s12124-018-9447-5>.

von Oertzen, C. (2017). Machineries of data power: Manual versus mechanical census compilation in nineteenth-century Europe. *Osiris* 32(1), 129–50.

<https://doi.org/10.1086/693916>.

Recommended Citation

Editorial Team (2023). Epistemologies of data. *On Education. Journal for Research and Debate*, 6(18).

https://doi.org/10.17899/on_ed.2023.18.0



Do you want to comment on this article? Please send your reply to editors@oneducation.net. Replies will be processed like invited contributions. This means they will be assessed according to standard criteria of quality, relevance, and civility. Please make sure to follow editorial policies and formatting [guidelines](#).