

### Are social experiments being hyped (too much)?

Neuwinger, Malte

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Neuwinger, M. (2023). Are social experiments being hyped (too much)? *TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis / Journal for Technology Assessment in Theory and Practice*, 32(3), 22-27. <https://doi.org/10.14512/tatup.32.3.22>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

RESEARCH ARTICLE

# Are social experiments being hyped (too much)?

Malte Neuwinger\*,<sup>1</sup> 

22

**Abstract** • Social experiments, also known as randomized controlled trials, are the subject of contentious discussions, giving rise to buzzwords such as ‘credibility revolution,’ ‘experimenting society,’ ‘global lab,’ or ‘empire of truth.’ While using exaggeration to illustrate opportunities and risks may well be justified, this research article analyzes to what extent the present debate is characterized by excessive hype. It finds that the transformative potential of social experiments is greatly overestimated, a judgment that applies to the reasoning of both proponents and critics.

*‘It became a thing – in academia and outside organizations. And then it became controversial, which, in a sense, is even better.’*

Nobel laureate Esther Duflo  
about the rise of social experiments  
(quoted in Parker 2010)

## Werden Sozialexperimente (zu sehr) gehypt?

**Zusammenfassung** • Sozialexperimente, auch bekannt als randomisierte kontrollierte Studien, werden kontrovers diskutiert, etwa unter den Schlagworten ‚Revolution der Glaubwürdigkeit‘, ‚Experimentiergesellschaft‘, ‚globales Labor‘ oder ‚Imperium der Wahrheit‘. Obwohl Übertreibung zur Verdeutlichung von Chancen und Risiken durchaus gerechtfertigt sein kann, untersucht dieser Forschungsartikel inwiefern die aktuelle Diskussion durch einen übermäßigen Hype geprägt ist. Im Ergebnis wird festgestellt, dass das transformative Potenzial von Sozialexperimenten weit überschätzt wird. Diese Diagnose gilt gleichermaßen für die Argumente von Unterstützern und Kritikern.

**Keywords** • hype, social experiment, RCT, instrument constituencies, tools-to-theories heuristic

This article is part of the Special topic “Technology hype: Dealing with bold expectations and overpromising” edited by J. Bareis, M. Roßmann and F. Bordignon. <https://doi.org/10.14512/tatup.32.3.10>

## The debate

There is little doubt that modern societies should use their best knowledge to improve people’s lives. Yet what exactly constitutes our best knowledge and how exactly it should be used are controversial questions. Over the past twenty years, a wave of ‘evidence-based policy making’ has provided one answer: Public policy should be tested through social experiments. Much like in clinical trials, such experiments randomly assign people to ‘treatment’ and ‘control’ groups: In the simplest case, one receives the new program, the other does not, and if the former program leads to better effects than the latter we may conclude that our program ‘works’. And much like in clinical trials, such an approach is seen as applying the organized skepticism of science to policy making: If we are uncomfortable with taking untested drugs, why would we be comfortable with subjecting ourselves to untested public policies?

This reasoning has given rise to a wave of excitement. As the titles of several recent books and articles inform us, ‘randomized controlled trials’ (RCTs), as social experiments have been labelled to fit the medical model, constitute a ‘credibility revolution’ in social scientific research (Angrist and Pischke 2010), the beginning of a ‘twenty-first century experimenting society’ (White 2019) in which ‘radical researchers are changing our world’ (Leigh 2018). Long-time commitment to the claim that RCTs will ‘revolutionize social policy’, as they revolutionized medicine (Duflo and Kremer 2005, p. 228), has earned three economists the 2019 Nobel Prize. (As explained below, the strong connection between medicine and social science seems overdrawn. Because it was the standard terminology before the current hype, I therefore speak of social experiments instead of

\* Corresponding author: [malte.neuwinger@uni-bielefeld.de](mailto:malte.neuwinger@uni-bielefeld.de)

<sup>1</sup> Faculty of Sociology, Bielefeld University, Bielefeld, DE



RCTs (Greenberg et al. 1999).) Among other things, social experiments have been used to check whether microcredits can help people rise out of poverty, whether disbursing money for relocation improves people's health and income, and whether a universal basic income has positive economic and psychological effects. While most prominent in Anglophone countries, social experiments are increasingly spreading around the world. Even countries lagging behind the trend, like Germany, have now begun to express interest (Faust 2020).

For some time now, the amount of attention social experiments receive has raised suspicions of hype. Critics have worried that many claims about experiments' superiority are epistemically unjustified, politically tendentious, and ethically questionable, because they involve unwarranted generalizations beyond the particular experimental context, promote small problem fixes at the expense of larger socio-economic effects, and are cavalier about people's rights (Kvangraven 2019; Picciotto 2012; Ravallion 2009). Pushing these critiques further, some speak of the emergence of a new 'empire of truth' that crowds out democratic deliberation through technocratic governance (Kelly and McGoey 2018) or an elitist 'global lab' that reduces people to the equivalent of test animals (Fejerskov 2022). According to critics, by combining scientific credibility, a strong media profile, and the support of philanthropic foundations supporters have turned social experimentation into a profitable 'scientific business model' (Bédécarrats et al. 2019, p. 750).

What should one make of this debate, whose strong rhetoric and self-perceived societal relevance is likely to baffle the uninitiated? I argue that proponents and critics have usefully raised awareness about the potentials and risks of social experiments, but by now the most radical factions of the debate are in danger of losing touch with reality. Drawing on open-ended interviews with twenty influential advocates, implementers, and funders, the main section of this article shows that practitioners are much more pragmatic about social experiments than academic discussions would make one believe. These interviews were conducted in 2022 and 2023, mostly via video call, targeting key social experiment supporters such as the International Initiative for Impact Evaluation (3ie) and the Abdul Latif Jameel Poverty Action Lab (J-PAL). In addition, conceptualizing the phenomenon of 'hype', the remainder of the article argues that social experiments should be seen as a tool that actively shapes the thinking of those involved in the debate. In particular, this tool promotes the misleading assumption that social experiments can easily be compared with drug trials and narrows down attention to a policy's 'impact' in the sense of causality.

## Hype and criti-hype

Debates about new scientific and technological developments often make it hard to tell which side is right. One reason for this is that both the relevant facts and the relevant criteria of 'right' and 'wrong' are part of the discussion. Another is that situations in

which things are still unfolding make it harder to identify vested interests and analytical blind spots. To mitigate this problem, technology assessment (TA) and especially science and technology studies (STS) recommend the principles of 'symmetry' and 'reflexivity'. Symmetry prompts researchers to 'maintain a posture of balanced skepticism' toward both sides of the debate. And reflexivity, as Stephen Hilgartner, Sheila Jasanoff and Hilton Simmet note in a 'living document' that circulates within the STS community, appeals to researchers' capability to 'become aware of the assumptions underlying their knowledge claims and, where necessary, to address specific blind spots and sources of bias or error'. Because things in the making involve what STS scholars call 'interpretive flexibility', one should rather be careful with strong epistemic or normative judgments.

My brief survey of the social experiment debate suggests that it is pretty much the opposite of symmetric or reflexive. Rather, 'credibility revolutions', 'experimenting societies', 'global labs', and 'empires of truth' are examples of hype in the technical sense of the term. While all science needs to go slightly beyond existing evidence to make useful inferences, 'inappropriate exaggeration' willingly sacrifices reasonable prediction in favor of generating excitement and enthusiasm. Indeed, though 'enthusiasm' may seem like an odd description of worries about technocracy and unethical human experimentation, the concept of hype applies equally to optimistic and pessimistic exaggeration. After all, both have the effect of impeding a clear assessment of the issue at hand (Intemann 2022, pp. 180–182). If not in 'content', radical critiques of social experiments therefore seem closely connected to their intellectual counterparts in 'form': While the normative evaluation flips from celebratory to alarmist, both sides agree that the implications for policy making will be radical.

Social experiments are not alone in being hyped both ways. Indeed, fighting positive overclaiming through negative overclaiming seems to be part of a general phenomenon, one that has been termed 'criti-hype' (Vinsel 2021). Many scholars have worried that widely hailed technologies like genetic engineering, nanotechnology, or social media will bring about a social dystopia. While such worries are quite reasonable in principle, the problem with criti-hypes is that they seem less interested in getting a handle on the problems identified than in imagining grim 'technoscientific futures' (Vinsel 2021). Ironically, then, the asymmetrical, unreflexive social experiment debate may feed not one but two 'scientific business models': one that acquires money and publicity through experimentation, and the other that does the same by criticizing that experimentation – and both do this though it is quite unclear whether social experiments will have the predicted effects one way or another.

## Will social experiments transform policy making?

Obviously, not all discussion of social experiments constitutes hype. But what does? As Intemann (2022) stresses, particularly

STS scholars have been rather vague about the criteria by which they identify hype. As indicated, her solution is to focus on claims that are both exaggerated and inappropriate. In other words, identifying hype involves *empirical assessments* as well as reasoned and explicit *value judgments*. What does this imply for the social experiment debate?

Let's begin with value judgments. Intemann suggests that relevant judgments may be divided into two parts, namely (1) the most important goals of communication and (2) the acceptable risk of getting things wrong: What should advocates and critics try to communicate and how bad would it be if their claims turned out to be false? In my judgment, the social experiment debate's goal should be to accurately communicate the potentials and risks of social experiments as tools of political decision making. If one gives at least some credence and weight to the worries of both sides of the debate – as I think one should – the risk of getting things wrong suggests a tension. Overstating the benefits and understating the risks might lead to increased global injustices caused by social experimentation, while the converse error might perpetuate badly informed and at worst harmful government policies at the expense of better ones.

Based on these value judgments (which may be disputed, but to me seem quite modest), one might already conclude that many commentators have indeed exaggerated inappropriately. Clearly, they have not even tried to give a balanced account of benefits and risks. On the other hand, one might argue that some exaggerations may nevertheless be useful because they clarify that diverging value judgments are premised on very different normative concerns: harm through experimentation vs. harm through business as usual (Parkhurst 2017, pp. 7–8). Here the second part of identifying inappropriate exaggeration comes in: empirical assessments. As I will show, both sides of the social experiment debate often exaggerate inappropriately because they invite unwarranted inferences about social experiments *given the evidence we have available* – namely that they predict transformative change on the basis of very limited facts. This becomes clear when considering that most practitioners appear to have adopted a 'new middle ground' between the 'well-rehearsed and polarized positions' of hyperts and criti-hyperts (Giselquist and Niño-Zarazúa 2015, p. 2).

One claim that pervades recent discussions is that evidence from social experiments is the 'gold standard' of evidence while anything else supposedly 'has no legitimacy and basis in reality' (Fejerskov 2022, p. 172; Gerber et al. 2014). Surprisingly, among people whose job consists in implementing and funding social experiments, very few seem to share this view. Instead, one employee of the Abdul Latif Jameel Poverty Action Lab (J-PAL), a major social experiment advocate, describes them as 'a tool in

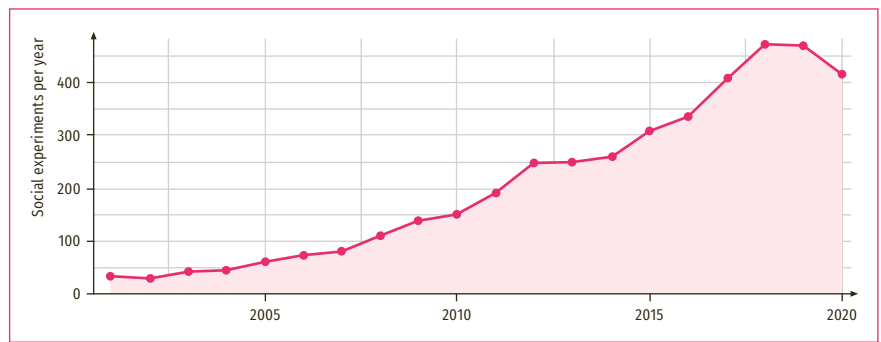


Fig. 1: Annual number of social experiments in international development, conducted between 2000–2020.

Source: Jie (2023)

the toolbox rather than the answer to all questions' (interview J-PAL 3). Another predicts that social experiments 'will become less [...] glorified. And they will become just a tool in the arsenal of governments' (interview J-PAL 2). An analyst at Arnold Ventures, a major philanthropic funder, worries about a 'point, and maybe we're here now, where there's just a lot of frustration with how often unsatisfying the answers to these questions are', sometimes because experiments produce no measurable effects, sometimes because interventions are badly implemented – and often because trying to tell these options apart is like 'banging your head against the wall' (interview Arnold Ventures).<sup>1</sup> Senior officials at the Behavioral Economics Team of the Australian Government similarly oscillate between confidence in the benefits of experiments and disillusionment (Ball and Head 2021, pp. 113–115). And so do researchers at the Behavioural Insights Team (BIT) in the UK, stressing that 'RCTs are not the answer to everything – you need to combine them with all kinds of other approaches' (interview BIT).

An accurate assessment of social experiments' transformative potential also needs a good grasp of how large the 'business' actually is and how much potential it has for growth. This is surprisingly difficult. For international development, one of the largest and best-documented fields, a few hundred social experiments are conducted every year, with a clear upward trend over time – compare Figure 1 to the current annual number of medical trials, which is around 60,000 (WHO 2023). In Northern countries, experiments in education are certainly on the rise, while other fields are smaller and precise numbers are harder to come by. (Much depends on the exact definition of a 'social experiment': How many people should it involve? How much does the 'treatment' need to differ from the status quo?) But even given clearer definitions and better data, the question is: Would, say, several thousand social experiments a year be enough to transform policy making?

<sup>1</sup> In the natural sciences, a similar problem is known as the 'experimenter's regress': Well-conducted experiments discover new facts, but these 'facts' only are facts if we know that the experiment was 'well-conducted', which we only 'know' because of the new 'facts' (Godin and Gingras 2002).

Part of the answer depends on whether governments are interested enough in the results to use them in their daily decision making. This issue has given practitioners some headaches – even strong evidence loses out to political strategy – but there are certainly ongoing efforts to establish collaborations with governments (Taddese 2021). Another part of the answer depends on whether the current business model is as profitable and sustainable as hypers and criti-hypers seem to believe. For international development, the assumption is that social experimentation will keep flourishing because ‘demand is twin-engined, driven by both the donor community and the academic world’ while ‘supply is largely shaped by a brand of scientific businesses and entrepreneurs’ (Bédécarrats et al. 2019, p. 750).

tend to be people-centered: ‘Hype cycle’ models suggest that people pass through different stages of excitement and frustration before an innovation finally brings productive development (Dedehayir and Steinert 2016). The reverse model is tool-centered: Rather than a passive thing people are more or less excited about, innovations can develop ‘a life of their own’ in the sense that (1) new tools prompt people to popularize them because their users come to materially depend on them, and (2) acquaintance with new tools shapes people’s thinking. While not explicitly framed as a theory of hype, the active role of innovative tools has been proposed as an important factor for developments in both science and politics. The ‘tools-to-theories heuristic’ suggests that the rising familiarity with statistical concepts

## *Isn’t talk of ‘impacts’, ‘outcomes’, and ‘results’ simply a symptom of rampant managerialism in public policy?*

As far as I can tell, however, this analysis overlooks a host of misaligned incentives among the actors being involved.

One primary misalignment is that most academics are interested in policy evaluation only if they can publish their results in academic journals. Almost all practitioners interviewed report a fundamental tension between academically interesting and practically relevant experimental work. The compromise is often to test small ‘nudges’ that are easy to implement and quick to evaluate (White 2014, pp. 21–22). Unfortunately, academically clever but tiny interventions are rarely useful for government policy – with the possible exception of behavioral science applications in governance, where the incentives of academics and public partners roughly converge on light touch interventions (Fels 2022). As one researcher at the German Institute for Development Evaluation (DEval) remarks, ‘even in the English-speaking world, it’s individual cases where it has really been win-win for both sides, where there’s been an academic publication and it also helped on the practical side’ (interview DEval). Practitioners also worry that funding may die down because experiments are too expensive, or that governments may lose interest because results take too long to become available. Overall, while social experiments are a significant phenomenon with benefits and risks, their transformative potential seems limited given the available evidence and the incentives of relevant actors.

### **Thinking through tools: how social experiments shape hypers’ and criti-hypers’ reasoning**

Having discussed whether social experiments will have a transformative effect on policy making (probably not), it is interesting to turn around and ask whether experiments may have an effect on hypers and criti-hypers themselves. Descriptions of hype

influenced theories of psychology (Gigerenzer 1991), while the concept of ‘instrument constituencies’ suggests that acquaintance with the notion of citizen panels affected prevailing thinking about political representation (Simons and Voß 2018). Social experiments seem to influence the thinking of both hypers and criti-hypers in a similar way, leading them to equate social experiments with drug trials and think of ‘impact’ in the narrow sense of causality.

While they fundamentally disagree about the implications, hypers and criti-hypers are united in comparing social experiments to clinical trials. To one side, the success of evidence-based medicine renders social experiments the obvious solution to social problems (Leigh 2018, chapter 2). To the other side, the fact that drug trials are increasingly outsourced to the Global South serves as an effective warning against the social inequalities scientific and technological innovations can produce (Fegerskov 2022, chapter 5). Such dissensus in consensus resembles the nineteenth century practice of ‘bundling’ loosely associated issues into large-scale, urgent, and contentious questions (the Eastern Question, the Jewish Question, etc.): Both sides of the debate merge social and medical experiments into a single Experimental Question, seeking to ‘raise the profile of their questions in order to draw attention to preferred solutions’ (Case 2018, p. 4). By thinking ‘through’ the tool, all experiments become the same – and depending on one’s inclinations more of them signify either a move toward ‘science’ or ‘technocracy’. Perhaps subconsciously, bundling medical and social experiments wins everyone involved argumentative mileage.

The odd feature of this apparent agreement is not only that hypers and criti-hypers rarely investigate the comparability of social and medical experiments explicitly, but that neither side seems to notice that a huge part of high-tech, high-stakes medicine is rarely subject to experimental evaluation, including surgery (Bothwell and Jones 2021). And this is despite the fact that

surgery is in many ways the better comparison. For instance, in both social and surgical interventions placebos and double blinding are much harder to implement than for drugs (for practical and ethical reasons) and their success crucially depends on the skills and motivations of those who implement them (respectively, surgical teams and public administrators).

The second way in which the tool of social experiments influences the thinking of hypers and criti-hypers alike is that it focuses their attention on the notion of ‘impact’ (Breslau 1997). Indeed, social experiments zoom in on a very particular interpretation of impact in the sense of the *precise estimation of causal effects, everything else being equal*. The logic of experiments – the point of which is to create a counterfactual ‘ceteris paribus world’ through randomization – provides the discussion with a definition of what impact *is*. Hypers and criti-hypers naturally disagree whether social experiments are the adequate tool to assess impact-as-causality (Bédécarrats et al. 2019; Gerber et al. 2014), but they rarely recognize that their understanding is in direct competition with the very different conception of impact as the *overall effect a policy has, considering the complexity of the real world*. The latter has long been the official definition of the OECD Development Assistance Committee (2021), which has never settled on experimental methods (Faust 2020, pp. 74–75). Evidently, thinking ‘through’ the tool of social experiments discourages advocates and critics from engaging other conceptions of impact.

It is possible to disregard such confusions as another oddity. Isn’t talk of ‘impacts’, ‘outcomes’, and ‘results’ simply a symptom of rampant managerialism in public policy? Yes, but not ‘simply’. After all, regardless of preferred vernacular it is difficult to deny that both the question (1) whether an observed improvement can be causally attributed to a newly implemented policy and (2) whether that policy really creates improvements in the grand scheme of things are highly important. Clear thinking about both (and their interconnection) is central to any policy debate and disregarding them as ‘neoliberal’ is just as unhelpful as trading off one against the other. In sum, while the social experiment debate has clarified the potential risks and benefits of social experimentation itself, narrow tool-based thinking seems to prevent greater clarity about what it means to ‘have a positive impact’ on people’s lives.

## Conclusion

Hypes put an interesting twist on the truism that ‘extraordinary claims require extraordinary evidence’. As extraordinary claims about the transformative potential of social experiments become the norm, the more cautious view that such claims are inappropriate exaggerations is put in the position to provide extraordinary evidence. Thus, people who have built a reputation around praising or criticizing social experiments will predictably not be convinced of my argument. And they should not be, because the evidence I have presented is indeed partial. It is possible that I

have misjudged the facts, that my interviewees have deceived me, or that my analysis is faulty in some other way. Still, I hope my argument can shift the burden of proof at least a little.

Critics might also ask whether less hype would actually lead to better decisions. Isn’t my argument based on a ‘technocratic fallacy’ (as one reviewer of this article opines): that ‘facts [...] are only an input [to policy making, my addition] that can (and should!) be ignored in the name of ideology’? It is certainly true that democratic politics cannot do without value judgments (just like the identification of hype) and that the selection of relevant facts requires deliberation. The process is complicated. Still, I agree with Parkhurst (2017) that good policy requires both values *and* facts. Just like insisting on the importance of impact is not *per se* neoliberal, insisting on the importance of facts is not *per se* technocratic.

**Funding** • This work received no external funding.

**Competing interests** • The author declares no competing interests.

## Research data

Interviewees consented that the information they provided can be used in publications, but without transcripts being open to the public. The Special topic editors have had access to the transcripts quoted in the article.

## References

- 3ie – International Initiative for Impact Evaluation (2023): 3ie development evidence portal. Available online at <https://developmentevidence.3ieimpact.org>, last accessed on 14. 08. 2023.
- Angrist, Joshua; Pischke, Jörn-Steffen (2010): The credibility revolution in empirical economics. How better research design is taking the con out of econometrics. In: *Journal of Economic Perspectives* 24 (2), pp. 3–30. <https://doi.org/10.1257/jep.24.2.3>
- Ball, Sarah; Head, Brian (2021): Behavioural insights teams in practice. *Nudge missions and methods on trial*. In: *Policy & Politics* 49 (1), pp. 105–120. <https://doi.org/10.1332/030557320X15840777045205>
- Bédécarrats, Florent; Guérin, Isabelle; Roubaud, François (2019): All that glitters is not gold. The political economy of randomized evaluations in development. In: *Development and Change* 50 (3), pp. 735–762. <https://doi.org/10.1111/dech.12378>
- Bothwell, Laura; Jones, David (2021): Innovation and tribulation in the history of randomized controlled trials in surgery. In: *Annals of Surgery* 274 (6), pp. e616–e624. <https://doi.org/10.1097/SLA.0000000000003631>
- Breslau, Daniel (1997): The political power of research methods. Knowledge regimes in U.S. labor-market policy. In: *Theory and Society* 26 (6), pp. 869–902. <https://doi.org/10.1023/A:1006802628349>
- Case, Holly (2018): *The age of questions. Or, a first attempt at an aggregate history of the eastern, social, woman, American, Jewish, Polish, bullion, tuberculosis, and many other questions over the nineteenth century, and beyond*. Princeton, NJ: Princeton University Press. <https://doi.org/10.23943/princeton/9780691131153.001.0001>
- Dedehayir, Ozgur; Steinert, Martin (2016): The hype cycle model. A review and future directions. In: *Technological Forecasting & Social Change* 108, pp. 28–41. <https://doi.org/10.1016/j.techfore.2016.04.005>

- Duflo, Esther; Kremer, Michael (2005): Use of randomization in the evaluation of development effectiveness. In: George Pitman, Osvaldo Feinstein and Gregory Ingram (eds.): Evaluating development effectiveness. World Bank series on evaluation and development, vol. 7. Piscataway, NJ: Transaction Publishers, pp. 205–231.
- Faust, Jörg (2020): Rigorose Wirkungsevaluierung. Genese, Debatte und Nutzung in der Entwicklungszusammenarbeit. In: *dms – der moderne staat – Zeitschrift für Public Policy, Recht und Management* 13 (1), pp. 61–80. <https://doi.org/10.3224/dms.v13i1.08>
- Fejerskov, Adam (2022): The global lab. Inequality, technology, and the experimental movement. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198870272.001.0001>
- Fels, Katja (2022): Who nudges whom? Expert opinions on behavioural field experiments with public partners. In: *Behavioural Public Policy*, pp. 1–37. <https://doi.org/10.1017/bpp.2022.14>
- Gerber, Alan; Green, Donald; Kaplan, Edward (2014): The illusion of learning from observational research. In: Dawn Langan Teele (ed.): *Field experiments and their critics. Essays on the uses and abuses of experimentation in the social sciences*. New Haven, CT: Yale University Press, pp. 9–32. <https://doi.org/10.12987/9780300199307-003>
- Gigerenzer, Gerd (1991): From tools to theories. A heuristic of discovery in cognitive psychology. In: *Psychological Review* 98 (2), pp. 254–267. <https://doi.org/10.1037/0033-295X.98.2.254>
- Gisselquist, Rachel; Niño-Zarazúa, Miguel (2015): What can experiments tell us about how to improve government performance? In: *Journal of Globalization and Development* 6 (1), pp. 1–45. <https://doi.org/10.1515/jgd-2014-0011>
- Godin, Benoît; Gingras, Yves (2002): The experimenters' regress. From skepticism to argumentation. In: *Studies in History and Philosophy of Science Part A* 33 (1), pp. 133–148. [https://doi.org/10.1016/S0039-3681\(01\)00032-2](https://doi.org/10.1016/S0039-3681(01)00032-2)
- Greenberg, David; Shroder, Mark; Onstott, Matthew (1999): The social experiment market. In: *Journal of Economic Perspectives* 13 (3), pp. 157–172. <https://doi.org/10.1257/jep.13.3.157>
- Intemann, Kristen (2022): Understanding the problem of 'hype'. Exaggeration, values, and trust in science. In: *Canadian Journal of Philosophy* 52 (3), pp. 279–294. <https://doi.org/10.1017/can.2020.45>
- Kelly, Ann; McGoey, Linsey (2018): Facts, power and global evidence. A new empire of truth. In: *Economy and Society* 47 (1), pp. 1–26. <https://doi.org/10.1080/03085147.2018.1457261>
- Kvangraven, Ingrid (2019): Impoverished economics? Unpacking the economics Nobel Prize. In: *openDemocracy*, 18. 09. 2019. Available online at <https://www.opendemocracy.net/en/oureconomy/impoverished-economics-unpacking-economics-nobel-prize>, last accessed on 10. 10. 2023.
- Leigh, Andrew (2018): *Randomistas. How radical researchers are changing our world*. New Haven, CT: Yale University Press. <https://doi.org/10.12987/9780300240115>
- OECD – Organization for Economic Co-operation and Development (2021): *Applying evaluation criteria thoughtfully*. Paris: OECD. <https://doi.org/10.1787/543e84ed-en>
- Parkhurst, Justin (2017): *The politics of evidence. From evidence-based policy to the good governance of evidence*. London: Routledge. <https://doi.org/10.4324/9781315675008>
- Parker, Ian (2010): The Poverty Lab. In: *The New Yorker*, 17. 05. 2010. Available online at <https://www.newyorker.com/magazine/2010/05/17/the-poverty-lab>, last accessed on 10. 10. 2023.
- Picciotto, Robert (2012): Experimentalism and development evaluation. Will the bubble burst? In: *Evaluation* 18 (2), pp. 213–229. <https://doi.org/10.1177/1356389012440915>
- Ravallion, Martin (2009): Should the randomistas rule? In: *The Economists' Voice* 6 (2), pp. 1–5. <https://doi.org/10.2202/1553-3832.1368>
- Simons, Arno; Voß, Jan-Peter (2018): The concept of instrument constituencies. Accounting for dynamics and practices of knowing governance. In: *Policy and Society* 37 (1), pp. 14–35. <https://doi.org/10.1080/14494035.2017.1375248>
- Taddese, Abeba (2021): *Meeting policymakers where they are. Evidence-to-policy and practice partnership models*. Washington, DC: Center for Global Development. Available online at <https://www.cgdev.org/sites/default/files/meeting-policymakers-where-they-are-background-paper.pdf>, last accessed on 10. 10. 2023.
- Vinsel, Lee (2021): You're doing it wrong. Notes on criticism and technology hype. In: *STS-News*, 01. 02. 2021. Available online at <https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5>, last accessed on 10. 10. 2023.
- White, Howard (2014): Current challenges in impact evaluation. In: *The European Journal of Development Research* 26 (1), pp. 18–30. <https://doi.org/10.1057/ejdr.2013.45>
- White, Howard (2019): The twenty-first century experimenting society. The four waves of the evidence revolution. In: *Humanities & Social Sciences Communications* 5 (1), pp. 1–7. <https://doi.org/10.1057/s41599-019-0253-6>
- WHO – World Health Organization (2023): *Global observatory on health research and development. Number of clinical trials by year, location, disease, phase, age and sex of trial participants (1999–2022)*. Available online at <https://www.who.int/observatories/global-observatory-on-health-research-and-development/monitoring/number-of-trial-registrations-by-year-location-disease-and-phase-of-development>, last accessed on 10. 10. 2023.



### MALTE NEUWINGER

is a PhD student with the Research Training Group 'World Politics: The emergence of political arenas and modes of observation in world society' at Bielefeld University. Because he is currently writing a thesis on the rise of social experiments, Malte hopes the hype will continue for some time into the future.