

## Invited Letter Rejoinder

**Cite this article:** Starke G, De Clercq E, Borgwardt S, Elger BS (2021). Why educating for clinical machine learning still requires attention to history: a rejoinder to Gauld *et al*. *Psychological Medicine* **51**, 2512–2513. <https://doi.org/10.1017/S0033291720004766>

Received: 1 November 2020  
Accepted: 10 November 2020  
First published online: 14 December 2020

**Author for correspondence:**

Georg Starke, E-mail: [georg.starke@unibas.ch](mailto:georg.starke@unibas.ch)

# Why educating for clinical machine learning still requires attention to history: a rejoinder to Gauld *et al*

Georg Starke<sup>1</sup> , Eva De Clercq<sup>1</sup>, Stefan Borgwardt<sup>2,3</sup>  and Bernice Simone Elger<sup>1,4</sup>

<sup>1</sup>Institute for Biomedical Ethics, University of Basel, Basel, Switzerland; <sup>2</sup>Department of Psychiatry, University of Basel, Basel, Switzerland; <sup>3</sup>Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany and <sup>4</sup>University Center of Legal Medicine, University of Geneva, Geneva, Switzerland

We are very grateful that Christophe Gauld, Jean-Arthur Micoulaud-Franchi and Guillaume Dumas have added their valuable comment to our article (Starke, De Clercq, Borgwardt, & Elger, 2020). We fully agree with their response, highlighting the importance of an appropriate framework for educating young psychiatrists (Gauld, Micoulaud-Franchi, & Dumas, 2020). Indeed, basic knowledge about the fundamentals of computer science, cognitive neuroscience, computational psychiatry, clinical practice as well as ethics seems crucial for a successful and responsible implementation of machine learning (ML) in psychiatry. Similarly, we fully concur with them and others (Grote & Berens, 2020) that developing an appropriate epistemological framework will be crucial to advance the ethical debates surrounding AI in healthcare.

Still, expanding on the useful practical guide Dr Gauld and his colleagues have provided to develop a curriculum fit for educational purposes, we would like to draw further attention to the persistent importance of teaching history of psychiatry. While this is no new demand (Shorter, 2008), it may not have received enough attention in the context of psychiatric ML yet. Of course, we are aware that curricula run the danger of being overburdened in the context of ML, and agree with Gauld *et al.* (2020) and McCoy *et al.* (2020) that training should focus on fundamental concepts. However, education about the historical development and employment of psychiatric classifications should be considered part of these fundamental issues and will remain crucial to counter potential ethical, clinical and conceptual pitfalls of ML in psychiatry. Once more, the example of schizophrenia seems particularly well suited to highlight these challenges.

With a view to ethical questions, education about the historical ramifications surrounding the development of particular classificatory concepts helps to elucidate the fact that they are human-made. When developing and using diagnostic ML tools in psychiatry, this may help to stress their historical contingency as heuristic concepts, countering tendencies to reify the categories which a particular system has been trained to classify (Hyman, 2010). Furthermore, attention to historical atrocities and gross abuse of power in psychiatry, e.g. during the Nazi era, can serve as a cautionary tale in educative settings, raising awareness for ethical pitfalls today (Strous, 2007). In fact, some old ethical problems of psychiatry may return under new guise with ML-based systems. For example, it has been argued that in the USA during the 1960s and 1970s, the diagnosis of schizophrenia was disproportionately applied to African-Americans connected to the civil rights movement, on account of their alleged aggressive behaviour (Metzl, 2009). Given that even today there remain significant disparities between ethnic groups with regard to the diagnosis of schizophrenia (Gara, Minsky, Silverstein, Miskimen, & Strakowski, 2019), educative curricula should draw attention to such historical injustices, fostering particular attention to discrimination and biases potentially ingrained in ML-based systems.

For the current clinical practice of psychiatry, obtaining a historically informed view seems highly beneficial as well. In particular, historical education may promote clinical qualities that critics fear could fade into the background with the introduction of ML systems. For example, looking closely at the original conditions under which a specific concept was introduced may inspire close attention to clinical context. Again, the case of schizophrenia can serve to illustrate this. The term ‘schizophrenia’ was famously coined by the Swiss psychiatrist Eugen Bleuler in 1908, arguably in rejection of a Kraepelinian nosology based on prognosis (Maatz & Hoff, 2014). In turn, Bleuler has been read as an early proponent of a bio-psycho-social model of disease, aiming for an understanding of the disorder that integrates the underlying neurobiology with individual psychological and social aspects (Maatz, Hoff, & Angst, 2015). In a similar vein, recent research has highlighted the irreducible and subjective psychological nature of Bleuler’s so-called first-rank symptoms, stressing the importance of the individual, lived experiences of patients for his psychopathology (Moscarelli, 2020). With regard to ML systems, teaching about the historical origins of the concept of schizophrenia

may thus serve to avoid an overly simplified view of the disorder and stress their cumulative nature. In other words, recent advances in ML notwithstanding, psychiatry will need to keep paying close attention to the social conditions of disorders as well as the individual phenomenological perspectives of patients.

Finally, with a view to conceptual questions, attention to the history of psychiatric theory will also remain fundamental to the development and improvement of diagnostic categories. We fully agree with Dr Gauld and his colleagues that an appropriate framework of medical epistemology requires a ‘to-ing and fro-ing’ between philosophy and science. However, in line with contemporary philosophy of science, we also hold that this process needs to retain attention to historical detail, in the sense of integrated history and philosophy of science (Chang, 2004). Kenneth Kendler has sketched the consequences of such a historical approach with regard to the classification of schizophrenia, driven by a process of ‘epistemic iterations’ (Kendler, 2009). Attempts to redefine psychiatric classification based on ML may thus need to reflect upon their own historically contingent role in this evolutive process, so that psychiatric nosology may mature ‘historically from top-down essentialist views of our categories to bottom-up empirically defined entities that reflect with increasingly accuracy the world as we can best understand it’ (Kendler, 2009).

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Conflict of interest.** None.

## References

- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. New York, NY: Oxford University Press.
- Gara, M. A., Minsky, S., Silverstein, S. M., Miskimen, T., & Strakowski, S. M. (2019). A naturalistic study of racial disparities in diagnoses at an outpatient behavioral health clinic. *Psychiatric Services*, 70(2), 130–134. doi: 10.1176/appi.ps.201800223
- Gauld, C., Micoulaud-Franchi, J.-A., Dumas, G. (2020). Comment on Starke *et al.*: ‘Computing schizophrenia: Ethical challenges for machine learning in psychiatry’: From machine learning to student learning: Pedagogical challenges for psychiatry. *Psychological Medicine*, 1–3. doi: 10.1017/S0033291720003906
- Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, 46(3), 205–211. doi: 10.1136/medethics-2019-105586
- Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology*, 6, 155–179. doi: 10.1146/annurev.clinpsy.3.022806.091532
- Kendler, K. S. (2009). An historical framework for psychiatric nosology. *Psychological Medicine*, 39(12), 1935–1941. doi: 10.1017/S0033291709005753
- Maatz, A., & Hoff, P. (2014). The birth of schizophrenia or a very modern Bleuler: A close reading of Eugen Bleuler’s ‘Die Prognose der Dementia praecox’ and a re-consideration of his contribution to psychiatry. *History of Psychiatry*, 25(4), 431–440. doi: 10.1177/0957154X14546606
- Maatz, A., Hoff, P., & Angst, J. (2015). Eugen Bleuler’s schizophrenia – a modern perspective. *Dialogues in Clinical Neuroscience*, 17(1), 43–49.
- McCoy, L. G., Nagaraj, S., Morgado, F., Harish, V., Das, S., & Celi, L. A. (2020). What do medical students actually need to know about artificial intelligence? *NPJ Digital Medicine*, 3(1), 86. doi: 10.1038/s41746-020-0294-7
- Metzl, J. (2009). *The protest psychosis: How schizophrenia became a Black disease*. Boston: Beacon Press.
- Moscarelli, M. (2020). A major flaw in the diagnosis of schizophrenia: What happened to the Schneider’s first rank symptoms. *Psychological Medicine*, 50(9), 1409–1417. doi: 10.1017/S0033291720001816
- Shorter, E. (2008). History of psychiatry. *Current Opinion in Psychiatry*, 21(6), 593–597. doi: 10.1097/YCO.0b013e32830aba12
- Starke, G., De Clercq, E., Borgwardt, S., & Elger, B. S. (2020). Computing schizophrenia: Ethical challenges for machine learning in psychiatry. *Psychological Medicine*, 1–7. doi: 10.1017/S0033291720001683
- Strous, R. D. (2007). Psychiatry during the Nazi era: Ethical lessons for the modern professional. *Annals of General Psychiatry*, 6, 8. doi: 10.1186/1744-859X-6-8