The vexatious black box of psychiatry may be smaller because of this issue

Tom Grimes[1,*]

[1]Texas State University, USA

*Author for correspondence:
Email: grimes@txstate.edu

Editorial

Psychiatry is burdened by a much larger black box than several other medical practices. Some practitioners have chosen fields where mechanical interventions address a patient's disability and so the black box is small. But psychiatry is one of those medical specialties whose practitioners confront perplexing unknowns. Putman notes the frequent instances of treatment failures, but he also notes that a worse situation is often a mild to moderate treatment *success*. Abject failures, as Putman observes, lead to a Plan B, another approach that might work [1]. But "mild success" can mean no follow-up at all. Presumably, these outcomes derive from what is hidden in the black box.

In order to better manage these anomalies, Al-Khathami and his colleagues have suggested that the Saudi Arabian model of primary health care can help. They make a convincing case that best practices in medicine require a psychiatric component attached to each primary care visit as part of a comprehensive healthcare model. As Putman notes, people are living longer and thus they are acquiring and manifesting pathology in ways that past generations – and clinicians – have not had to manage or have had the opportunity to observe and learn from. Al-Khathami's and colleagues' recommendations offer both a way to treat more people in pain, but also give clinicians more data to study, more clinical experiences to apply, and therefore increase the chance that the unknowable in psychiatry may become less impenetrable [2].

Carillo's commentary may be more important than he realizes. He notes an unfortunate attribute that characterizes the practice of medicine across all specialties: The overwhelming deluge of information that confronts physicians. For understandable reasons, new, originally collected data are given preference by science and medical journals as opposed to contextualizing commentaries and reflective analyses of previously published studies. However, this molecularization of knowledge – hundreds of individual studies, scattered across time and publication venues – makes encompassing all the parts of a clinician's practice difficult. And so as Carillo observes, clinicians tend to retreat to the familiar, the well-practiced, the safe diagnostic or therapeutic choice. Carillo is advocating the use of a trademarked software program, the Umbrella Collaboration, which he describes as a "tertiary evidence [synthesizer]" to help psychiatrists manage new information more effectively [3]. The Umbrella Collaboration notwithstanding, Carillo highlights one of the most important issues in this volume. There is no efficient way to make use of emerging knowledge if it is suspended in an uncontextualized epistemological space. Clinicians' ability to mentally model (the consolidation of disparate information into a semantic whole that can be held in working memory) is central to sound medical practice. Although Carillo does not use the term mental modeling, his recommendation for the use of tertiary analysis can build mental models. Parsimonious summaries of complex, abstract

information help physicians take in new information in a way they can apply it. Carillo is correct in pushing back against scientists who look unfavorably at reviews and meta-analyses. It is true that the synthesization of tertiary evidence rounds corners, generalizes, and deletes detail. But that is not the problem. The problem is the careless management of generalizations and summaries. In the hands of a competent collator (presumably for Carillo it is the Umbrella Collaboration), a summary or meta-analysis will improve the quality of information physicians receive. Think about it: Why should summaries and meta-analyses be of less quality than primary research? The answer to that question brings us to the Grimes and Sussman commentary.

Grimes and Sussman suggest that error is no less likely in original data studies than in meta-analyses. Their case study centers on commercially produced media violence, which researchers in this specialty have believed, *ipso facto*, is injurious to all consumers [4]. This belief put the study of media violence on the wrong track at its inception. As Grimes and Sussman document, media violence researchers for 45 years went about creating methodological techniques that assumed that commercial media violence harms everyone. Thus, the task became one of measuring the effects of that putative harm rather than questioning if that harm exists in the first place.

Media violence research, at least the line of inquiry that assumed all commercial media violence is psychopathologically harmful to anyone who consumes it, is no longer fashionable. Recent research has offered a more plausible hypothesis: Harm afflicts those predispositioned by psychopathology [5,6]. And in any event, social media and artificial intelligence are the new villains. The upshot is that Carillo's advocacy of summaries and meta-analyses can help address the torrent of information that paradoxically adds to the effect of the black box rather than ameliorating it.

Vatan and Lester assert that "empirical studies have shown that hopelessness-based *stress proneness patterns* predict both depression symptoms and a diagnosis of depression." Therefore, they advocate the use of a three-dimensional scale that assesses one's vulnerability to the effects of hopelessness, helplessness, and haplessness. Vatan and Lester, like many diagnosticians, are using the best available tools to deal with abstract, almost evanescent psychiatric conditions, at least in their less severe forms [7]. As Servaas and colleagues note [8], the more ill a person is, the more scales successfully align with the parameters that characterize ill people. Consequently, the scales Vatan and Lester are promoting may help patients who need the help the most.

The black box of psychiatry may soon be less of a problem. Putman proposes a black box resolution by suggesting that when a clinician develops a theory of patient illness that is deductively derived, it is more difficult for that clinician to consider alternative hypotheses because the act of deduction sets the analytical strategy on a singular pathway [1]. This, of course, is exacerbated by Carillo's point that there is too much information for physicians to process, information that *could* make alternative explanations more likely if that new information could be absorbed [3]. Putman advocates holding a diagnosis in suspension and then applying inductive reasoning by using the evidence that is available. Interestingly, Putman's suggestion, and the physicist Neils Bohr's suggestion [9] that human decision-making may follow a pattern similar to the superpositioning effect in quantum mechanics, originate from the same perspective. Bohr suggested that decision-making is not a linear, deductive process but rather one that exists in multiple states simultaneously – a superpositioned state. Physicist Eugene Wigner, in 1960, elaborated on Bohr's idea by arguing that the mathematics of quantum superpositioning can be applied to help explain how people make decisions [10]. Khrennikov and colleagues, publishing in the journal *Nature* in 2018, expanded the concept by asserting that decision-making is not a linear function, rather, "the key point is that the processing of information by the brain involves superpositions of [decision] states" such that "[superpositioning] guarantees exponentially fast [processing]" by connecting with other neural networks simultaneously rather than following linear neural pathways [11]. Superpositioning produces a probability wave that describes more likely, and less likely, decision outcomes similar to that of a Gaussian distribution where the peak collects the most likely set of decisions [12]. When a person makes a decision, the wave collapses at a single decision point. When Putman's clinician makes a decision after first holding in abeyance all possible alternatives, the "wave collapses" and settles on a diagnosis. Anything that inhibits the creation of that decision wave, such as rigidly deductive reasoning, may eliminate viable alternative diagnoses.

Large Language Models (LLMs), a category of artificial intelligence, have the potential to replicate the superpositioning characteristic of human decision-making. Several scholars have speculated that LLMs will eventually simulate cognition's superpositioning function because an LLM has the potential to integrate lexical, verbal, and visual indicators of illness in its database *and* execute a superpositioning function in assessing that database so a diagnosis can be made, or a therapy can be prescribed [13,14]. This could ameliorate some of the problems Putman and Carillo have noted in their commentaries as well as make Vatan's and Lester's hopelessness, helplessness, and haplessness construct easier to diagnose. As of now, superpositioning by LLMs is not possible due to hardware and software limitations [15]. But there are researchers [16–20] who are exploring ways to apply quantum wave theory to psychiatric diagnoses. This application, once the hardware is developed, may eventually be the key to diminishing the size of psychiatry's black box.

References

1. Putman III HP. Clinical reasoning, medical error, and treatment failure. Curr Res Psychiatry. 2025;5(1):9–16.

2. Al-Khathami AD, Alharbi LS, Alomari SA, Alqahtani AA, Alfadhli DS. Integrating mental healthcare into primary healthcare services. Curr Res Psychiatry. 2025;5(1):1–8.

3. Carrillo B. From complexity to clarity: The Umbrella Collaboration® and the future of tertiary evidence synthesis in psychiatry. Curr Res Psychiatry. 2025;5(1):22–25.

4. Grimes T, Sussman K. A commentary on methodological considerations for studying the psychological impact of social media. Curr Res Psychiatry. 2025;5(1):17–21.

5. Markey PM, Ferguson CJ. Moral Combat: Why the War on Video Games is Wrong. Dallas, Texas: BenBella Books, Inc.; 2017.

6. Grimes T, Lasser J. The importance of the null hypothesis in the formulation of theory in media psychology. New Ideas in Psychology. 2025 Apr 1;77:101142.

7. Vatan S, Lester D. The relationships between hopelessness, helplessness, haplessness and their effects on psychological well being. Curr Res Psychiatry. 2025;5(1):26–30.

8. Servaas MN, Schoevers RA, Bringmann LF, van Tol MJ, Riese H. Trapped: rigidity in psychiatric disorders. Lancet Psychiatry. 2021 Dec;8(12):1022–4.

9. Bohr N. The Quantum Postulate and the Recent Development of Atomic Theory1. Nature 1928;121:580–90.

10. Wigner EP. Remarks on the mind-body question. In: Mehra J, Editor. Philosophical reflections and syntheses. Berlin, Heidelberg: Springer Berlin Heidelberg; 1995. pp. 247–60.

11. Khrennikov A, Basieva I, Pothos EM, Yamato I. Quantum probability in decision making from quantum information representation of neuronal states. Sci Rep. 2018 Nov 1;8(1):16225.

12. Gu B, Desai RJ, Lin KL,Yang J. Probabilistic medical predictions of large language models. Npj: Digital Medicine. 2024;7(367).

13. Aerts D, Aerta Arguëlles J, Beltran L, Geriente S, de Bianchi MS, Sozzo S, et al. Quantum entanglement in physical and cognitive systems: A conceptual analysis and general representation. The European Physical Journal Plus. Eur Phys J Plus. 2019;134:493.

14. Wang B, Li Q, Melucci M, Song D. Semantic Hilbert space for text representation learning. In: The World Wide Web Conference. 2019, May. pp. 3293–9.

15. Kim J, Podlasek A, Shidara K, Feng L, Ahmed A, Bernardo D. Limitations of large language models in clinical problem-solving arising from inflexible reasoning. Scientific Reports. 2025;15:39426.

16. Tešić J, Stewart NK, Grimes T, Sokal JO. An adaptation of Hilbert spaces and Cauchy sequences to serve as the basis for the development of a Large Language Model that can diagnose depression. 2026. Unpublished manuscript.

17. Obradovich N, Khalsa SS, Khan WU, Suh J, Perlis RH, Ajilore O, *et al*. Opportunities and risks of large language models in psychiatry. NPP—Digital Psychiatry Neuroscience. 2024;2:8.

18. Sadeghi M, Richer R, Egger B, Schindler-Gmelch L, Rupp LH, Rahimi F, Berking M, Eskofier BM. Harnessing multimodal approaches for depression detection using large language models and facial expressions. Npj Ment Health Res. 2024 Dec 23;3(1):66.

19. Xin L, McDuff D. Evaluating and Enhancing Probabilistic Reasoning in Language Models. Google Health; 2024. Available from: https://research.google/blog/evaluating-and-enhancing-probabilistic-reasoning-in-language-models/.

20. Binz M, Akata E, Bethge M, Brändle F, Callaway F, Coda-Forno J, et al. A foundation model to predict and capture human cognition. Nature. 2025 Aug;644(8078):1002–9.