Can we train better medical intuition? Exploring the potential of debiasing interventions

Nina Franiatte[1,2], Wim De Neys[1], Esther Boissin[3]

[1]Université Paris Cité, LaPsyDÉ, CNRS, 46 rue Saint Jacques, 75005 Paris, France

[2]Research and Development Team, Onepoint, 2 rue Marc Sangnier, 33110 Bègles, France

[3]Department of Psychology, Cornell University, Ithaca, USA

Abstract

Human judgment is often prone to biases, and healthcare professionals are no exception. In clinical environments – characterized by high pressure, time constraints, and information overload – intuitive impressions can sometimes override statistical reasoning, leading to severe consequences such as diagnostic errors. There is an urgent need to identify effective strategies for reducing clinical decision-making errors and improving patient safety. In this paper, we examine the roots of logical fallacies and present promising debiasing interventions designed to mitigate such biases. Especially, we review debiasing training procedures that have been shown to significantly boost logico-mathematical reasoning, producing durable improvements across reasoning tasks and populations. Critically, these improvements occur as early as the initial intuitive stage, allowing faster and more accurate responses – which is particularly relevant to clinical contexts. We also discuss recent alternative procedures that help identify the conditions under which training is most effective. Overall, these findings suggest that short debiasing training can cultivate reliable intuitive judgments, offering a scalable, ecologically valid path to reducing medical decision-making errors.

Keywords: Medical decision-making · Healthcare professionals· Reasoning · Intuition · Debiasing training

Fast thinking leads us astray? Exploring the roots of logical fallacies

Each year, around 5% of American adults experience a diagnostic failure in medicine, leaving nearly 800,000 Americans dead or permanently disabled (Newman-Toker et al., 2024). Such failures are partly due to the short duration of clinical visits and the limited time available for experts' reflection (Topol, 2024). This striking statistic highlights how intuitive, fast reasoning can result in dramatic consequences in high-stakes contexts. It also underscores the importance of research on human decision-making, and the critical need to identify ways of reducing biased reasoning to improve patient safety (Croskerry & Nimmo, 2011; Reyna, 2008).

Reasoning research has long shown that human judgment is prone to biases (e.g., Kahneman, 2011; Tversky & Kahneman, 1974); healthcare professionals are no exception. Croskerry (2003) documented a wide range of biases that affect clinical reasoning, particularly in emergency settings, where access to patients' information (e.g., medical history) is limited and time for decision-making is restricted. The availability heuristic is one such example: Physicians may overestimate the likelihood of recently encountered diseases while underestimating those not seen for a long time, thereby increasing the risk of diagnostic errors (Croskerry, 2003). A related example is the base-rate neglect bias. During the COVID-19 pandemic, many people mistakenly believed vaccines to be ineffective because a majority of hospitalized patients were vaccinated. What was overlooked is that vaccinated individuals vastly outnumbered the unvaccinated in the population, making it statistically expected that many hospitalizations would occur among the larger group (i.e., among vaccinated people; see De Neys, 2023; Devis, 2021). These examples highlight how intuitive impressions can sometimes override statistical reasoning, and lead to systematic diagnostic errors.

Sparked in the 1970s by the pioneering work of Daniel Kahneman and Amos Tversky (e.g., Tversky & Kahneman, 1974), dual-process theory offers explanations for such reasoning errors and has since shaped thinking across various disciplines, including medicine (e.g., Croskerry, 2009; Djulbegovic et al., 2012). In this theory, human reasoning is described as an interplay between two types of processes or 'systems': A fast intuitive one (often called 'System 1'), and a slower, more effortful, deliberate one (often called 'System 2', e.g., Evans & Stanovich, 2013; Kahneman, 2011). The intuitive 'System 1' is typically conceived as the system that makes decisions based on heuristic cues by relying on prior world knowledge and beliefs, while the deliberate ‘System 2’ is typically conceived as the system responsible for effortful thinking, allowing people to reason logically and probabilistically.

One of the most influential viewpoints in the reasoning field – known as the default-interventionist view – posits that ‘System 1’ first produces an (often erroneous) intuitive response, which may then be monitored and corrected by the deliberate ‘System 2’ (Evans & Stanovich, 2013). In other words, ‘System 2’ must correct ‘System 1’ to reach the correct solution. However, given that reasoners tend to minimize demanding computations, they will often apply intuitive processes by default without considering that the correct answer might be different (Kahneman, 2011). Hence, people often remain biased.

Yet more recent evidence has nuanced this picture. Conflict detection studies have demonstrated that even when people give a biased intuitive response, they are not entirely blind to it and do show some sensitivity to their errors (e.g., De Neys et al., 2011; Pennycook et al., 2015). These studies typically contrast the response given on a conflict version of a problem – where heuristic intuitive answers conflict with logico-mathematical norms – to the no-conflict (control) version of the same problem, where heuristic intuitive answers align with normative principles. For instance, the infamous bat-and-ball problem, initially presented by Frederick (2005) illustrates this phenomenon: ‘*A bat and ball together cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?’*. This conflict version triggers a first intuitive response (i.e., 10 cents) that does not align with the correct logico-mathematical response (i.e., 5 cents). In contrast, in the no-conflict (control) version of the same problem, the critical relational ‘more than’ statement is removed, leading to similar intuitive and logico-mathematical answers (‘*A bat and ball together cost $1.10. The bat costs $1.00. How much does the ball cost?*’). Empirical evidence shows that responses to conflict versions of reasoning problems tend to be slower, associated with lower confidence, and higher electrodermal activity compared to matched no-conflict control problems (e.g., De Neys et al., 2011; Frey et al., 2016). Altogether, these findings indicate that reasoners possess at least some implicit sensitivity of the core underlying logical and probabilistic norms, even when they fail to apply them (e.g., De Neys, 2017).

Building on this idea, studies using a two-response paradigm (Thompson et al., 2011) have shown that even correct answers can sometimes be generated immediately at the intuitive stage, before deliberation intervenes (e.g., Bago & De Neys, 2017, 2019). In this paradigm, participants are asked to give two consecutive responses to each reasoning problem. First, they are required to give their initial intuitive response under time pressure while simultaneously performing a secondary memory task. Since deliberation requires time and cognitive resources, restricting both maximally enforces intuitive responding (Bago & De Neys, 2019). Immediately afterwards, they are presented with the problem again and can take all the time they need to think about it and give their final deliberate response. Surprisingly, two-response findings revealed that a non-neglectable proportion of correct answers appears already in the initial intuitive stage (e.g., Bago & De Neys, 2017, 2019; Newman et al., 2017; Raoelison et al., 2020, 2021). These findings challenge the assumption that logico-mathematical reasoning always requires 'System 2' intervention and have motivated the development of hybrid dual-process models (Bago et al., 2021; Baron & Gürçay, 2017; Białek & De Neys, 2017). According to these models, 'System 1' is not limited to the generation of erroneous intuitive responses. It can also produce logical intuitions, and 'System 2' is engaged selectively, typically when conflict detection signals the need for further reflection.

This raises a central question: Where do logical intuitions come from? Stanovich (2018) introduced the concept of mindware, referring to the rules, principles, and strategies acquired through learning and practice. With repeated exposure, logico-mathematical principles can be automatized such that they become directly accessible at the intuitive level (De Neys & Pennycook, 2019). In the absence of relevant mindware, intuition rarely aligns with normative standards; with sufficient practice, however, correct responses can be generated effortlessly. Evidence from medicine supports this account. Notably, it has been shown that novices and trainees tend to rely heavily on slow, deliberate reasoning, whereas experienced clinicians – having practiced more extensively – often rely on fast intuitive judgments that are highly accurate (Croskerry & Nimmo, 2011). Such rapid intuitive judgments made by healthcare professionals are referred to as clinical intuitions – that is, early, subjective assessments formed before diagnostic information becomes available (Zélis et al., 2019). For instance, Zélis et al. (2019) asked nursing staff and physicians to rate the severity of illness and their level of concern immediately upon a patient's arrival, prior to history taking, physical examination, or access to any diagnostic results. Results showed that they could intuitively and reliably predict the likelihood that elderly patients admitted to emergency departments would die or experience other adverse outcomes within 30 days. Similarly, in a study by Brabrand et al.

(2014), the first physician to assess each patient was asked to estimate the subjective probability of in-hospital mortality on a scale from 0 to 100%. These predictions were made immediately after the initial evaluation, without waiting for test results or additional information. The study demonstrated that using only these clinical intuitions, healthcare professionals could accurately identify patients at increased risk of dying while admitted to their medical admission unit.

In sum, these findings show that intuitive reasoning is not necessarily prone to biases: When relevant principles have been sufficiently automatized, 'System 1' can yield accurate responses. In such cases, relying on the intuitive route is entirely appropriate and allocating 'System 2' resources would be a waste of time and energy. However, it is equally clear that correct intuitive responses remain fairly rare. In most empirical studies, biased responses dominate on conflict reasoning problems, and in clinical practice, diagnostic errors continue to represent a significant challenge. From both a theoretical and applied perspectives, the challenge is therefore not merely to document the existence of logical intuitions, but to identify how their frequency and reliability can be increased, particularly in high-stakes contexts such as medical decision-making.

Can we do anything about it? Overview of (our) debiasing work

This practical concern has fueled a long line of research on debiasing interventions, designed to reduce reasoning errors and promote more accurate judgments (Lilienfeld et al., 2009; Milkman et al., 2009). Early efforts in this area produced mixed outcomes. While some interventions showed promising results, many yielded only modest effects, leading some researchers to question whether robust debiasing was feasible at all (e.g., Evans et al., 1994; Fischhoff, 1982). More recently, however, a new wave of studies has provided stronger evidence that short, targeted training interventions can significantly improve logico-mathematical reasoning (Boissin et al., 2021, 2022, 2023b, 2024b; Claidière et al., 2017; Franiatte et al., 2024a, 2024b, 2025; Hoover & Healy, 2017; Isler et al., 2020; Morewedge et al., 2015; Purcell et al., 2020; Trouche et al., 2014).

Notably, in their work, Boissin and Franiatte have applied a typical training design involving three phases: A pre-test, an intervention block, and a post-test (see Figures 1A and 1B). During pre- and post-tests, participants solve reasoning problems presented in the two-response format, enabling researchers to distinguish intuitive 'System 1' from deliberative 'System 2' responses. In the intervention block, participants in the training group receive short text explanations[1] highlighting both the common biased response and the correct reasoning procedure. For example, in the bat-and-ball problem, the explanation makes explicit why the intuitive answer of 10 cents is incorrect and demonstrates the correct calculation (ball = 5 cents; bat = $1.05). In contrast, participants in the control group solve the same problems without any feedback (Figures 1A and 1B).

The results of such interventions are striking. For instance, in Franiatte et al. (2024a), around 40% of trained participants improved their performance after receiving a short training about the bat-and-ball task. Critically, these debiasing effects were robust across tasks (e.g., conjunction fallacy, base-rate neglect; Boissin et al., 2022) and persisted for at least two months (Boissin et al., 2022; Franiatte et al., 2024a). Collectively, this body of work suggests that even brief interventions can significantly improve reasoning performance.

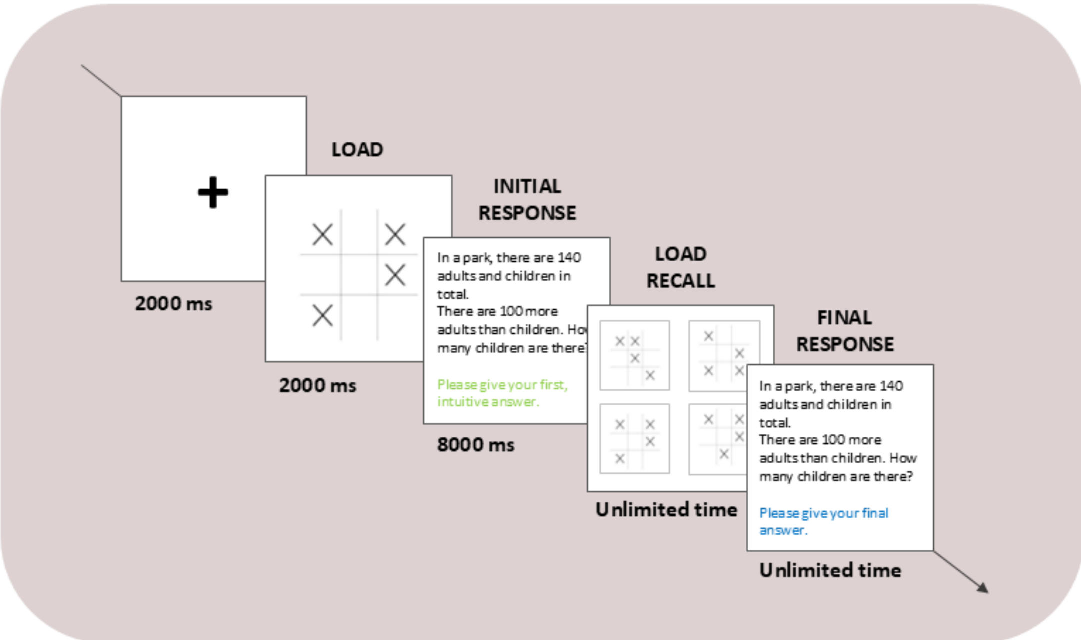

Figure 1A. Example of a typical bat-and-ball trial using the two-response paradigm (Thompson et al., 2011).

[1] The interested reader can try these interventions themselves on the following website: https://cogitum-site.powerappsportals.com/homePage

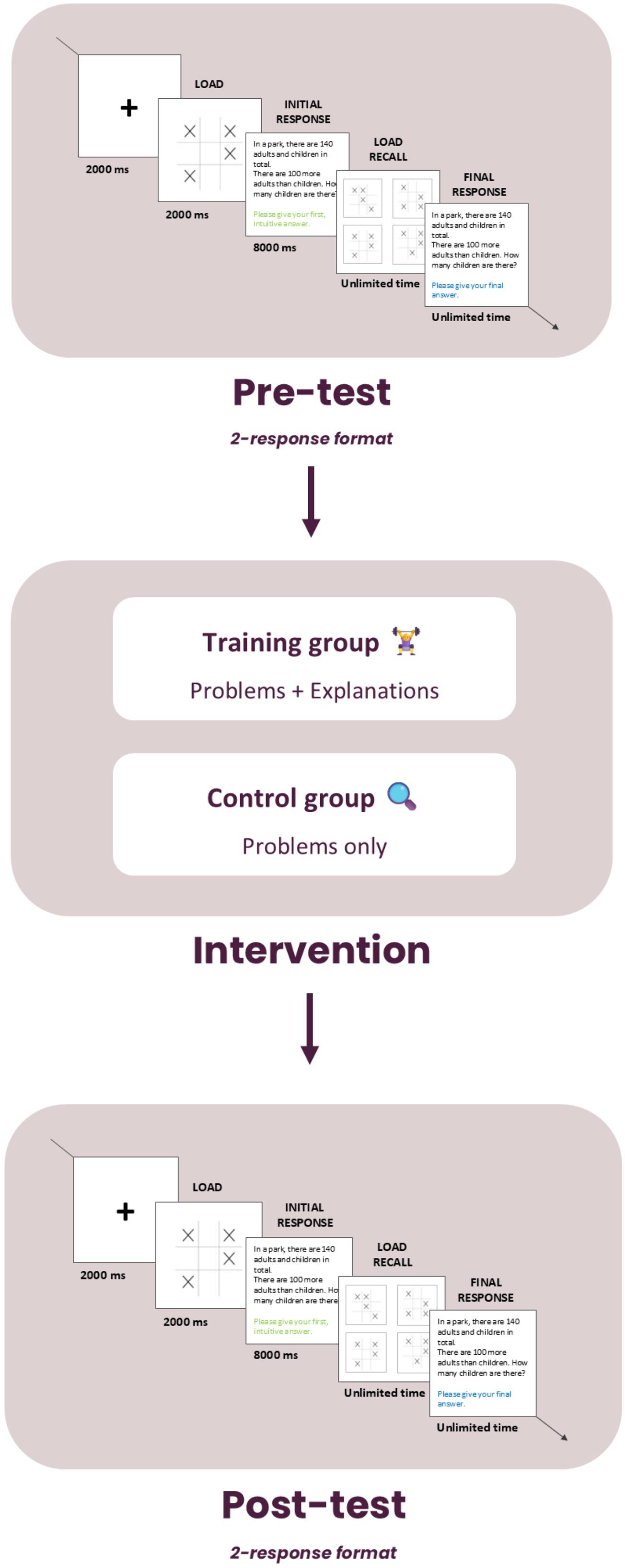

Figure 1B. Example of a typical debiasing training design involving three phases: A pre-test, an intervention block, and a post-test. During pre- and post-tests, reasoning problems are presented using the two-response paradigm (Thompson et al., 2011).

A key theoretical question concerns the nature of the training effect: Does training make participants more likely to deliberate, detecting and overriding their erroneous intuitions? Or does it instead lead them to intuitively apply the correct solution strategy without the need for correction? The traditional assumption is that training boosts 'System 2' processes, shifting reasoners from biased 'System 1' answers to deliberate 'System 2' corrections (e.g., Kahneman, 2011). However, strategies focused on promoting deliberation have repeatedly shown limited effectiveness (e.g., Lawson et al., 2020; Milkman et al., 2009). Deliberation is itself fallible – errors can stem from poor reasoning strategies (Pennycook et al., 2015) – and people often lack the time or motivation to engage in deliberate thinking (Boissin et al., 2021). Hence, even if such interventions succeed, they do so by optimizing a fundamentally suboptimal route. Deliberation is slow, resource-intensive, and in clinical contexts it is often unavailable because of high pressure, time constraints, and cognitive load. Thus, improving reasoning by relying more heavily on 'System 2' processes is unlikely to scale effectively to real-world emergency medical decision-making. By contrast, interventions that reshape 'System 1' processes – so that correct responses are generated intuitively – target the faster and more efficient pathway. In medical practice, if training enables clinicians to intuitively generate correct responses, it could offer a more powerful and ecologically valid path to reducing diagnostic errors. From this perspective, the central promise of debiasing is not simply to strengthen error correction, but to train intuition itself so that accurate responses emerge automatically from the outset.

Using the two-response paradigm, a consistent pattern has emerged across multiple studies: After debiasing training, improvements are observed primarily at the initial intuitive stage, rather than during subsequent deliberation stage. This suggests that training reshapes intuition directly, rather than merely teaching participants to correct their errors after the fact. A clear illustration comes from the bat-and-ball problem: Before training, we observed only about 15% of correct responses. After receiving a brief text explanation, this proportion increased to around 47% in the post-test (Boissin et al., 2021). Importantly, this improvement was not simply due to participants reflecting more in the second stage of the task: The bulk of the gains occurred immediately, at the very first response stage. By highlighting the relevance of the logical principles, the training intervention boosts the activation strength of correct intuitions so that they outcompete conflicting incorrect intuitions. Consequently, after training, the correct response can be generated automatically, without the need for further deliberation (Boissin, 2023). In other words, once trained, participants' intuitions shifted from yielding the common biased answer (i.e., '10 cents') to producing the correct response (i.e., '5 cents') automatically.

Moreover, these effects proved robust across contexts and tasks. In follow-up work, similar improvements were observed on other canonical reasoning problems, such as base-rate neglect and the conjunction fallacy (Boissin et al., 2022). Importantly, the benefits of training sustained, as correct intuitive responding remained elevated for at least two months after training (Boissin et al., 2021, 2022; Franiatte et al., 2024a). They also proved stable across different populations and languages. For example, interventions conducted with French-speaking participants produced comparable results to those obtained in English-speaking samples (Franiatte et al., 2024b), suggesting that the effects are neither language-specific nor culturally restricted.

Taken together, these findings strongly support what has been termed the 'trained intuitor' view. According to this perspective, training does not simply encourage people to engage 'System 2' more often; rather, it boosts 'System 1' so that the correct solution strategy is applied quickly and effortlessly. This represents a qualitative shift in how accuracy is achieved: Reasoners are no longer dependent on slow, effortful correction, but can rely on fast, effortless intuitions that are already aligned with normative principles.

Critically, recent findings suggest that debiasing interventions are most effective when certain favorable conditions are met. One central factor is the availability of relevant mindware – that is, prior knowledge of logico-mathematical principles that can be reactivated and applied intuitively (see above; Boissin, 2021; Stanovich, 2018). Against this backdrop, it becomes clear that training explanations do not create new logical concepts out of nothing; instead, they strengthen knowledge that is already present but not yet sufficiently automatized to dominate the competing heuristic.

Consequently, it is important to emphasize that debiasing interventions are not miraculous. A brief text explanation cannot create new logical knowledge from scratch. Rather, training is thought to operate by reactivating existing logical intuitions that were too weak to dominate the competing heuristic response. Supporting this view, Boissin et al. (2024b) found that training was markedly less effective among the Himba population in rural Namibia, who have limited access to informal education compared to those living in an urban environment or attending school. Likewise, developmental work by Raoelison et al. (2021) showed that older children – who had received more formal instruction in mathematics and probability – were more likely than younger children to generate correct intuitive responses. Together, these findings support the idea that logical intuitions emerge gradually as principles are acquired and automatized, and that debiasing effectiveness depends on the availability of such prior knowledge.

However, even among individuals with comparable, rich educational backgrounds, some consistently benefit from debiasing while others remain biased. For instance, after a bat-and-ball debiasing training, Boissin et al. (2021) still observed 32% of biased reasoners throughout their study (i.e., reasoners who gave a majority of incorrect intuitive and deliberate responses). Recent work started to tackle this issue by providing alternative procedures designed to boost the training efficacy. For instance, Franiatte et al. (2024a) showed that repeated test sessions can significantly boost intuitive reasoning performance compared to a single test session.

Likewise, Franiatte et al. (2025) tested a video-based debiasing approach and found that it benefited a larger group of reasoners and produced better long-term retention effects than the text-based training approach. Hence, to date, different forms of training have been applied to debias the intuitions of as many reasoners as possible, depending on the context and the specific biases to be addressed.

Additionally, prior knowledge and extensive repetition alone do not fully account for variability in training outcomes. Recent studies suggest that thinking dispositions – such as open-mindedness, overconfidence, or receptivity to pseudo-profound statements – play a critical role in shaping debiasing success. Boissin and Pennycook (2025) showed that individuals who failed to benefit from training tended to be less open-minded and more overconfident in their initial judgments. Thus, debiasing effectiveness appears to hinge on at least two complementary factors: The availability of automatized knowledge that can be reactivated, and the cognitive dispositions that determine whether individuals are willing and able to reconsider their first impressions.

Indeed, available evidence also suggests that even if training might not have succeeded in getting all biased people to reason more accurately, it might have helped them better detect their incorrect answers. In line with previous findings, biased reasoners in Franiatte et al. (2024a) showed increased response doubt when they erred on conflict problems after training, and this doubt was stronger among those who later became more accurate. These effects were more pronounced with repeated training, suggesting that increased doubt may serve as a precursor to the intervention effect. Overall, this highlights the role of metacognitive monitoring in reasoning and suggests that repeated training can also enhance these processes. These considerations carry important potential implications for medicine. Clinical practice routinely places physicians in conditions where time, stress, and information overload restrict deliberation. In such environments, relying on 'System 2' correction is rarely feasible. Training programs that can reshape 'System 1' processing – so that accurate responses arise automatically – offer a much more realistic route to improving decision-making. For instance, short explanatory interventions could be adapted into continuing medical education modules, case-based learning, or digital simulations (e.g., Poupard et al., 2025). Rather than encouraging clinicians to slow down in situations where slowing down is impossible, the goal would be to strengthen the accuracy of the fast route that clinicians often inevitably have to rely on.

A crucial next step is to determine how well laboratory-based training paradigms generalize to ecologically valid clinical contexts. To date, most debiasing work has relied on canonical reasoning tasks such as the bat-and-ball or base-rate neglect problems. Although these tasks capture fundamental cognitive mechanisms, medical decision-making involves richer and more ambiguous cues. Future studies need to test whether similar training formats – short explanations, repeated exposure, video-based interventions – can successfully reduce diagnostic errors in realistic clinical scenarios. Embedding debiasing modules into simulation-based medical training, for example, would allow researchers to measure whether trained intuitions improve triage accuracy, risk assessment, or diagnostic calibration under time pressure.

More broadly, the challenge is not to suppress intuition but to differentiate between reliable and unreliable intuitions. The hybrid dual-process perspective emphasizes that 'System 1' is not inherently biased; it can produce both heuristic errors and logical insights. For medicine, this means that interventions should aim not at replacing intuition with deliberation per se, but at cultivating the kinds of intuitions that align with probabilistic and logical principles. If clinical training succeeds in making these logical responses the default intuitive output, physicians will be better equipped to act swiftly and accurately when deliberation is not an option.

The potential impact is substantial. Each year, millions of patients are affected by diagnostic errors. Even a modest increase in the frequency of accurate intuitive responses could translate into dramatic improvements in patient outcomes and healthcare efficiency. The evidence reviewed here suggests that such improvements are possible: With the right training, intuitive reasoning can be reshaped into a powerful ally rather than a source of bias. The challenge ahead lies in adapting and scaling these interventions to medical contexts. Doing so could help bridge the gap between cognitive science and clinical practice – transforming insights about human reasoning into tangible advances for patient safety.

References

Aczel, B., Szollosi, A., & Bago, B. (2016). Lax monitoring versus logical intuition: The determinants of confidence in conjunction fallacy. *Thinking and Reasoning*, *22*(1), 99–117. https://doi.org/10.1080/13546783.2015.1062801

Bago, B., Bonnefon, J.-F., & De Neys, W. (2021). Intuition Rather Than Deliberation Determines Selfish and Prosocial Choices. *Journal of Experimental Psychology: General, 150*(6), 1081. https://doi.org/10.1037/xge0000968

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109. https://doi.org/10.1016/j.cognition.2016.10.014

Bago, B., & De Neys, W. (2019). The Smart System 1: evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, *25*(3), 257–299. https://doi.org/10.1080/13546783.2018.1507949

Baron, J., & Gürçay, B. (2017). A meta-analysis of response-time tests of the sequential two systems model of moral judgment. *Memory & Cognition, 45*(4), 566–575. https://doi.org/10.3758/s13421-016- 0686-8

Białek, M., & De Neys, W. (2017). Dual processes and moral conflict: Evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, *12*(2), 148–167. https://doi.org/10.1017/s1930297500005696

Boissin, E., Caparos, S., & De Neys, W. (2023b). No easy fix for belief bias during syllogistic reasoning?. *Journal of Cognitive Psychology*, *35*(4), 401-421. https://doi.org/10.1080/20445911.2023.2181734

Boissin, E., Caparos, S., Raoelison, M., & De Neys, W. (2021). From bias to sound intuiting: Boosting correct intuitive reasoning. *Cognition*, *211*. https://doi.org/10.1016/j.cognition.2021.104645

Boissin, E., Caparos, S., Voudouri, A., & De Neys, W. (2022). Debiasing System 1: Training favours logical over stereotypical intuiting. *Judgment and Decision Making*, *17*(4), 646–690. https://doi.org/10.1017/S1930297500008895

Boissin, E., Josserand, M., De Neys, W., & Caparos, S. (2024b). Debiasing thinking among non-WEIRD reasoners. *Cognition*, *243*, 105681. https://doi.org/10.1016/j.cognition.2023.105681

Boissin, E., & Pennycook, G. (2025). Who benefits from debiasing?. *Cognition*, *262*, 106166. https://doi.org/10.1016/j.cognition.2025.106166

Brabrand, M., Hallas, J., & Knudsen, T. (2014). Nurses and physicians in a medical admission unit can accurately predict mortality of acutely admitted patients: a prospective cohort study. *PloS one*, *9*(7), e101739. https://doi.org/10.1371/journal.pone.0101739

Claidière, N., Trouche, E., & Mercier, H. (2017). Argumentation and the diffusion of counter-intuitive beliefs. *Journal of Experimental Psychology: General*, *146*(7), 1052–1066. https://doi.org/10.1037/xge0000323

Croskerry, P. (2003). The Importance of Cognitive Errors in Diagnosis and Strategies to Minimize Them. *Academic Medicine*, *78*(8), 775–780.

Croskerry, P. (2009). Clinical cognition and diagnostic error: applications of a dual process model of reasoning. *Advances in health sciences education*, *14*(Suppl 1), 27-35. https://doi.org/10.1007/s10459-009-9182-2

Croskerry, P., & Nimmo, G. R. (2011). Better clinical decision making and reducing diagnostic error. *Journal of the Royal College of Physicians of Edinburgh*, *41*(2), 155–162. https://doi.org/10.4997/JRCPE.2011.208

De Neys, W. (Ed.) (2017). Dual Process Theory 2.0. Oxon, UK: Routledge.

De Neys, W. (2023). Advancing theorizing about fast-and-slow thinking. *Behavioral and Brain Sciences*, *46*, e111. https://doi.org/10.1017/S0140525X2200142X

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0015954

De Neys, W., & Pennycook, G. (2019). Logic, fast and slow: Advances in dual-process theorizing. *Current Directions in Psychological Science*, *28*(5), 503–509. https://doi.org/10.1177/0963721419855658

Devis, D. (2021). Why are there so many vaccinated people in hospital?. Retrieved from https://cosmosmagazine.com/health/covid/why-are-there-so-many-vaccinated-people-inhospital/

Djulbegovic, B., Hozo, I., Beckstead, J., Tsalatsanis, A., & Pauker, S. G. (2012). Dual processing model of medical decision-making. *BMC medical informatics and Decision Making*, *12*(1), 94. https://doi.org/10.1186/1472-6947-12-94

Evans, J. S. B. T., Newstead, S. E., Allen, J. L., & Pollard, P. (1994). Debiasing by Instruction: The Case of Belief Bias. *European Journal of Cognitive Psychology*, *6*(3), 263–285. https://doi.org/10.1080/09541449408520148

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 422–444). chapter, Cambridge: Cambridge University Press

Franiatte, N., Boissin, E., Delmas, A., & De Neys, W. (2024a). Boosting debiasing: Impact of repeated training on reasoning. *Learning and Instruction*, *89*, 101845. https://doi.org/10.1016/j.learninstruc.2023.101845

Franiatte, N., Boissin, E., Delmas, A., & De Neys, W. (2024b). Adieu Bias: Debiasing Intuitions Among French Speakers. *Psychologica Belgica*, *64*(1), 42. https://doi.org/10.5334/pb.1260

Franiatte, N., Boissin, E., Delmas, A., & De Neys, W. (2025). Debiasing in motion: Boosting sound intuiting through animated video training. *Acta Psychologica, 258*. https://doi.org/10.1016/j.actpsy.2025.105131

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Hoover, J. D., & Healy, A. F. (2017). Algebraic reasoning and bat-and-ball problem variants: Solving isomorphic algebra first facilitates problem solving later. *Psychonomic Bulletin and Review*, *24*(6), 1922–1928. https://doi.org/10.3758/s13423-017-1241-8

Isler, O., Yilmaz, O., & Dogruyol, B. (2020). Activating reflective thinking with decision justification and debiasing training. *Judgment and Decision Making*, *15*(6), 926–938. https://doi.org/10.1017/s1930297500008147

Kahneman, D. (2011). Thinking, fast and slow. New York, NY: Farrar, Straus and Giroux.

Lawson, M. A., Larrick, R. P., & Soll, J. B. (2020). Comparing fast thinking and slow thinking: The relative benefits of interventions, individual differences, and inferential rules. *Judgment and Decision making*, *15*(5), 660-684. https://doi.org/10.1017/S1930297500007865

Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving Debiasing Away Can Psychological Research on Correcting Cognitive Errors Promote Human Welfare? *Perspectives on Psychological Science*, *4*(4), 390–398. https://doi.org/https://doi.org/10.1111/j.1745-6924.2009.01144.x

Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How Can Decision Making Be Improved? *Perspectives on Psychological Science*, *4*(4), 379–383. https://doi.org/https://doi.org/10.1111/j.1745-6924.2009.01142.x

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing Decisions: Improved Decision Making With a Single Training Intervention. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 129–140. https://doi.org/10.1177/2372732215600886

Newman, I. R., Gibb, M., & Thompson, V. A. (2017). Rule-based reasoning is fast and belief-based reasoning can be slow: Challenging current explanations of belief-bias and base-rate neglect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(7), 1154-1170. https://doi.org/10.1037/xlm0000372

Newman-Toker, D. E., Nassery, N., Schaffer, A. C., Yu-Moe, C. W., Clemens, G. D., Wang, Z., ... & Siegal, D. (2024). Burden of serious harms from diagnostic error in the USA. *BMJ Quality & Safety*, *33*(2), 109-120. http://dx.doi.org/10.1136/bmjqs-2021-014130

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). Everyday Consequences of Analytic Thinking. *Current Directions in Psychological Science*, *24*(6), 425–432. https://doi.org/10.1177/0963721415604610

Poupard, M., Larrue, F., Bertrand, M., Liguoro, D., Tricot, A., & Sauzéon, H. (2025). Using virtual reality for enhancing neuroanatomy learning by optimizing cognitive load and intrinsic motivation. *Computers & Education*, 105332. https://doi.org/10.1016/j.compedu.2025.105332

Purcell, Z. A., Wastell, C. A., & Sweller, N. (2020). Domain-specific experience and dual-process thinking. *Thinking & Reasoning, 27*(2), 239-267. https://doi.org/10.1080/13546783.2020.1793813

Raoelison, M., Keime, M., & De Neys, W. (2021). Think slow, then fast: Does repeated deliberation boost correct intuitive responding? *Memory & Cognition*, *49*, 873–883. https://doi.org/10.3758/s13421-021-01140-x

Raoelison, M., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*. https://doi.org/10.1016/j.cognition.2020.104381

Reyna, V. F. (2008). A theory of medical decision making and health: fuzzy trace theory. *Medical decision making*, *28*(6), 850-865. https://doi.org/10.1177/0272989X08327066

Stanovich, K. E. (2018). Miserliness in human cognition: the interaction of detection, override and mindware. *Thinking and Reasoning, 24*(4), 423–444. https://doi.org/10.1080/13546783.2018.1459314

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. https://doi.org/10.1016/j.cogpsych.2011.06.001

Topol, E.J. (2024). Toward the eradication of medical diagnostic errors. *Science, 383*(6681), eadn9602. https://doi.org/10.1126/science.adn9602

Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, *143*(5), 1958–1971. https://doi.org/10.1037/a0037099

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, *185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

Zelis, N., Mauritz, A. N., Kuijpers, L. I., Buijs, J., de Leeuw, P. W., & Stassen, P. M. (2019). Short-term mortality in older medical emergency patients can be predicted using clinical intuition: a prospective study. *PloS one*, *14*(1), e0208741. https://doi.org/10.1371/journal.pone.0208741