Journal of Education, 2025
Issue 99, http://journals.ukzn.ac.za/index.php/joe http://dx.doi.org/10.17159/2520-9868/i99a01

Probabilistic professional judgement in teaching

Wayne Hugo

School of Education, University of KwaZulu-Natal, Pietermaritzburg, South Africa
hugow@ukzn.ac.za
https://orcid.org/0000-0002-1464-3977

(Received: 12 December 2024; accepted: 30 April 2025)

Abstract

Subjective Bayesian reasoning offers a framework for understanding how teachers actively refine their professional judgement in response to the inherent uncertainties of the classroom. Drawing on Luhmann's systems theory, Simon's bounded rationality, and Shalem's work on professional knowledge, in this paper I demonstrate how Bayesian reasoning models how teachers navigate three fundamental challenges: the operational separation between teaching and learning systems (creating inherent unpredictability); cognitive limitations that necessitate satisficing solutions (efficient ways to predict and decide); and the systematic development of professional knowledge through academic and diagnostic classifications (building the basis for better predictions). Through constructed scenarios, I demonstrate how novice teachers often begin with fragile priors based on theoretical knowledge and personal experience, which undergo dramatic updates when confronted with the mismatch between expectations and classroom realities (significant prediction errors). As teachers gain experience, they develop more robust and refined priors—belief systems that can incorporate new evidence while maintaining stable overall patterns, reflecting increasingly sophisticated predictive models. This evolution reflects the development of diagnostic classifications that guide professional decision-making. I show that subjective Bayesian reasoning provides a formal mechanism for modelling belief updating in professional judgement. While teachers may not engage in explicit probabilistic calculations, I argue that subjective Bayesian reasoning underlies the development of fast and frugal heuristics that become increasingly expert predictive tools with experience. By integrating Bayesian reasoning with established theories of professional knowledge development, a theoretical framework is offered that uses probability formally to demonstrate how teachers learn to make effective decisions by managing the inherent uncertainties and constraints of classroom teaching.

Keywords: Bayesian reasoning, professional judgement, satisficing, uncertainty, teacher development, prediction error, robust priors

Introduction

Teaching is fundamentally an exercise in decision-making under uncertainty (Borko et al., 2008), stemming from both classroom complexity and the impossibility of directly controlling student learning, which necessitates predictive judgement. Every day, teachers make hundreds of choices about instruction, assessment, and classroom management without being able to predict their outcomes perfectly (Bishop & Whitfield, 1972). As teachers develop from novices to experts, they must learn to navigate this uncertainty through increasingly sophisticated professional judgement (Berliner, 1986, 1987).

The analysis integrates three theoretical perspectives that, when viewed through a Bayesian lens (Joyce, 2004), show how teachers develop expertise in decision-making. Bayesian probability—named after Thomas Bayes (1701–1761)—represents a shift from viewing probability as a long-run frequency of events to understanding it as a measure of an individual's degree of belief, formed and updated through experience and reasoning (Bayes, 1763). In this framework, teaching is conceptualized as a continuous process of belief revision in response to new classroom evidence.

Luhmann's systems theory (2002/2012) provides the first layer of analysis. It highlights why teaching inherently involves probability rather than certainty. Given the operational closure between teaching and learning systems, teachers cannot directly access learners' mental states. As a result, they must rely on probabilistic estimates of whether their interventions have been understood or internalized. This detachment between communication and cognition structures teaching as an act of inference rather than control.

Simon's (1957) theory of bounded rationality adds a second layer. It explains why teachers do not search exhaustively for optimal solutions but instead develop satisficing strategies—efficient, experience-based heuristics shaped by the complexity of classroom contexts and the cognitive limitations of real-time decision-making. These heuristics serve as compact predictive models that guide moment-to-moment teaching choices under uncertainty.

The third layer comes from Shalem's work (2014, 2017) on professional knowledge, which traces how teachers transform these heuristics gradually into more disciplined forms of judgement. As they integrate theoretical insight with experiential knowledge, teachers refine their predictive beliefs, shifting from reactive patterns to more principled interpretive engagement with practice and theory.

Seen together, these perspectives frame teaching as a process of continuous Bayesian updating. Early-career teachers often experience dramatic belief revisions in response to prediction errors—such as when classroom realities contradict pedagogical assumptions (Bobadilla-Suarez et al., 2022; Bertram, 2023). This often leads to a temporary dependence on externally structured tools like scripted lesson plans, which provide more predictable scaffolds during periods of high uncertainty. Over time, experienced teachers move beyond these scripts, developing adaptive and context-sensitive practices that reflect accumulated experience and more calibrated belief systems (Sawyer, 2001; Winch, 2017). I argue that teachers engage in continuous probabilistic inference about how their teaching strategies might influence learning outcomes (Shafto & Goodman, 2008). This process involves predicting likely effects, observing results, and updating beliefs—a naturally occurring form of Bayesian reasoning driven by the need to reduce predictive error and improve effectiveness through structured stages of professional growth. While teachers do not calculate probabilities explicitly, they develop increasingly sophisticated heuristics that approximate Bayesian updating (Gigerenzer et al., 1999), enabling them to make effective decisions under uncertainty. By formalizing this intuitive process within a Bayesian framework, we gain deeper insights into the probabilistic mechanisms underlying teacher development and expertise.

Through constructed scenarios I demonstrate how this theoretical framework helps explain common patterns in teacher development and offers vignettes for supporting teachers' growth from novice to expert decision-makers. These scenarios illustrate how teachers develop from relying on fragile prior beliefs (leading to unreliable predictions) to establishing robust priors that enable stable yet flexible professional practice (grounded in well-calibrated predictive models).

Literature review and theoretical foundations

Longitudinal development of teacher decision-making

Research demonstrates that teacher decision-making progresses loosely through three distinct phases. Sawyer's (2001) decade-long study tracking three teachers in The United States provides an empirical account of this developmental trajectory.

In the initial "survival" phase, teachers often prioritise reducing immediate uncertainty by relying on prescribed routines and scripted frameworks. Sawyer (2001) documented how Ellen, a mathematics teacher, focused on "keeping above water" by minimizing unpredictable situations, implementing "relatively prescriptive approaches to math, such as lecturing, explaining rules and taking formulaic approaches" (p. 46). Bertram's research (2023) using the work of the Initial Teacher Education Research Project (Deacon, 2016) confirmed this pattern, and noted the reality shock of novice teachers trying to apply the teaching strategies learned at universities, the incessant demand on their time, and their struggle with school discipline (Bertram, 2023). It seems that these novice teachers needed a basic level of predictability regarding these urgent requirements before they could focus on the substantive aspects of teaching such as planning worthwhile lessons and learning activities, preparing high quality assessments, and thinking deeply about how to make conceptual connections between and within topics (Bertram, 2023).

The second phase involves what Sawyer (2001) identified as "seeing what works for students" (p. 47), marked by increased experimentation such as testing actively hypotheses about different approaches like cooperative learning and technology integration. Teachers try out different methods to gather information and start to refine their predictive sense of what works for them and what does not.

The third phase manifests as "maturity," characterised by "a more explicit use of multiple and thematically integrated approaches" (Sawyer, 2001, p. 47). Teachers demonstrate nuanced contextual decision-making guided by a robust, yet flexible, predictive framework. Rather than dramatic shifts between approaches, experienced teachers exhibit what Sawyer calls "a measured approach" (p. 54), thoughtfully evaluating and selectively incorporating new ideas against their well-calibrated predictive models while maintaining effective practices.

Systems theory and educational uncertainty: Luhmann's perspective

Niklas Luhmann's work on education articulated a key source of uncertainty in teaching—the tension between causality and freedom (Qvortrup, 2024) where teachers attempt to cause certain learning outcomes without being able to control directly the actual learning process of a student, who is free to decide to learn or not. At the heart of his analysis is the concept of autopoietic systems—self-producing and self-maintaining systems that operate according to their own internal logic (Luhmann, 1986, 1984/1995). This concept fundamentally challenges traditional input-output models of education by positioning teaching and learning as distinct systems, creating an inherent gap that necessitates probabilistic judgement.

The principle of operational closure (Luhmann, 1992) means that systems operate based on their own internal structures, with students' learning systems ultimately controlling their own processes from within themselves. This operational closure is not absolute isolation; it describes how systems process external influences according to their own internal logic. As a result, teachers must operate probabilistically, estimating the likelihood that their teaching strategies will influence learning successfully.

This necessity gives rise to what Luhmann calls the attribution of causality—a concept emerging from the condition of "double contingency" (Baraldi et al., 2017, p. 75), where both teaching and learning systems face mutual uncertainty about each other's responses. This double contingency takes on special significance in education, where teachers must construct probability estimates based on situational observations and patterns rather than relying on direct cause-effect relationships (Qvortrup, 2024).

As Luhmann and Schorr (2000) argued, causality in education is both impossible and necessary—impossible given the closed nature of each student's cognitive system, but necessary because teachers must develop ways of working that produce learning effects reliably. This paradox is addressed through what Luhmann (1984/1995, 2002/2012) termed structural coupling, where teaching probabilistically influences or irritates the learning system rather than controlling it (Qvortrup, 2024). This coupling provides a pathway for communication—the basic operation of social systems—while acknowledging that specific outcomes cannot be dictated.

In this way, Luhmann's theory demonstrates why probabilistic reasoning must be at the heart of teaching. Teachers must continually interpret and re-evaluate the likely success of different strategies, using past observations to update their probability estimates—an ongoing process akin to Bayesian updating, driven by the need to manage the inherent uncertainty of educational situations while bridging the gap between their need to act and students' freedom to respond unpredictably.

Bounded rationality in teaching: Simon's framework

Herbert Simon's theory of bounded rationality, developed in *Models of Man* (1957), fundamentally challenged fully rational assumptions about human decision-making. His critique targeted the *homo economicus* model, which presumed that decision-makers possess perfect information and unlimited processing capacity. Instead, Simon proposed that human cognitive limitations make such optimization unrealistic, forcing reliance on more efficient ways to predict and decide, an insight particularly relevant for understanding teacher decision-making (Lee & Porter, 1990).

Simon introduced the concept of satisficing to explain how people operate under constraints. Rather than seeking optimal solutions, individuals settle for options that meet a threshold of good enough predictive success (Simon, 1956, 1957). This framework helps explain why educators, faced with complex classroom situations, adopt practical, workable solutions rather than theoretically optimal ones. In teaching, this means that educators set acceptable levels for student engagement or learning outcomes and choose instructional strategies that meet these criteria, rather than searching for an elusive optimal method.

To address these limitations further, Simon (1976) introduced procedural rationality in acknowledging that decision-making processes themselves must be efficient within cognitive constraints. People use heuristics—mental shortcuts—that help them respond adaptively to complex, uncertain situations. These heuristics can be viewed as simplified rules that efficiently approximate Bayesian updating, allowing teachers to adjust their beliefs and predictions quickly based on new classroom experiences without complex calculations.

Simon's work extended beyond explaining bounded rationality to explore how expertise develops through pattern recognition. Studying chess masters, he found that experts do not necessarily consider more information than novices. Instead, they recognise familiar patterns and retrieve associated predictive responses from memory. As he noted, "The situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition" (Simon, 1992, p. 155) triggering a well-honed prediction.

This pattern recognition capacity is central to understanding the difference between novice and expert decision-makers. Research by Chase and Simon (1973) on chess players illustrated that masters could recall meaningful game positions almost perfectly because of these patterns, whereas novices struggled, especially with positions lacking strategic coherence. However, if the chess pieces were randomly placed on the board, this recall advantage dwindled away. Expert decision-making emerges not from expanded cognitive resources but from sophisticated heuristics encoded as patterns in long-term memory that enable rapid and effective prediction.

Through pattern recognition, experts develop richer sets of prior experiences that allow more accurate predictions about outcomes. This enhanced ability to anticipate and interpret classroom situations means that experts effectively update their beliefs in a Bayesian manner.

This development path shows that while bounded rationality limits perfect optimisation, experts' pattern recognition enables them to transcend some limitations through procedurally rational heuristics that are able to project forward probabilities of specific actions. The framework complements Bayesian perspectives by explaining how teachers make satisficing decisions, while developing increasingly sophisticated pattern recognition skills through experience, that are used to make probabilistic predictions about which decision will produce which results under given conditions.

Professional judgement in teaching: Shalem's framework

Professional judgement emerges as a key mediating factor in teaching, bridging the theoretical challenges identified by Luhmann's systems theory (1984/1995, 2002/2012) and Simon's concept of bounded rationality (1957). Teachers face a dual challenge: they must influence student learning without direct control over it while making decisions under cognitive constraints and with incomplete information. Shalem and Slonimsky (2010), Slonimsky and Shalem (2006), and De Clercq and Slonimsky (2014) have provided a theoretical framework for understanding how teachers develop the professional judgement necessary to navigate these challenges by building better predictive models of practice.

De Clercq and Shalem (2014) distinguished between two types of pedagogical content knowledge (PCK) that develop through epistemological labour (the intellectual work of refining one's understanding and practice). The first type, PCK1, involves organising teaching over time—sequencing and pacing content, using coherent lesson structures, establishing routines, and designing learning activities focused on structuring instruction to manage complexity and increase predictability. The second type, PCK2, encompasses specialised knowledge about how to enable learners to understand the meanings, rules, and procedures of the subject matter, providing them with "epistemological access" to new knowledge (requiring more nuanced predictions about student thinking) (De Clercq & Shalem, 2014, p. 140; see, too, Morrow, 1994).

Wally Morrow's pilot analogy illustrates this distinction.

> A good teacher of piloting has in the back of his mind an understanding of what is involved in flying an aeroplane . . . You can contrast such a teacher . . . who understands what the bigger thing is . . . from one who follows a book which says in lesson one you need to do this and in lesson two you need to do that and it is never properly tied together. (cited in Shalem & Slonimsky, 2010, p. 21)

Developing from PCK1 to PCK2 involves four key processes of epistemological labour: distantiation; appropriation; research; and articulation (Shalem & Slonimsky, 2006)—processes that actively refine the teacher's internal predictive model. Distantiation allows teachers to step back and examine their assumptions critically, moving beyond immediate practices and establishing cognitive distance from taken-for-granted knowledge. Appropriation involves integrating theoretical knowledge into existing understanding, adapting it for specific contexts—essentially, making the strange familiar. Research enables teachers to systematically investigate the effects of their teaching practices systematically, while articulation involves communicating and making public their knowledge and judgements. Through these processes, teachers engage in activities akin to Bayesian updating.

They adjust their prior beliefs about effective teaching strategies based on new theoretical insights and practical evidence, improving their predictive accuracy over time.

Relying solely on PCK1 limits teachers' ability to adapt to unexpected situations since procedural routines may not address the deeper understanding required for effective teaching. As De Clercq and Shalem (2014) have noted, scripted lesson plans can incorporate routines of PCK1 but fail to support the development of PCK2, which is essential for learners' epistemological access. Appropriate judgement about how close or far a learner is from what is correct depends on understanding the disciplinary rules—a key aspect of PCK2, giving what the bigger thing is about teaching (enabling better prediction of student learning needs).

Shalem (2014) further developed her framework through engagement with Abbott's (1988) work on professional knowledge. Abbott's distinction between academic and diagnostic classifications offered important insights into how predictive judgement develops. Academic classifications are organised frameworks that establish boundaries within disciplinary knowledge, providing theoretical concepts like the "zone of proximal development" and "epistemological access" that help teachers understand the bigger thing beyond procedural routines (Shalem, 2014, p. 98) and form structured prior hypotheses about learning. Diagnostic classifications, in contrast, are practice-oriented and they guide decision-making in specific cases. They enable professionals to interpret and act in particular situations by highlighting relevant features for predicting outcomes.

By integrating academic and diagnostic classifications, teachers develop a comprehensive predictive framework that informs their expectations and interpretations. This integration supports more accurate belief updating since theoretical knowledge provides a basis for forming priors, while practical experience supplies the evidence needed to calibrate likelihood estimates.

The progression from PCK1 to PCK2 represents a fundamental challenge in teaching practice, especially in light of Luhmann's observation that teaching and learning operate as separate systems (Luhmann & Schorr, 2000), and Simon's account of bounded rationality, which highlights the cognitive and contextual constraints that shape teacher decision-making (Simon, 1957). To bridge this gap, teachers must engage in systematic epistemological labour that develops both academic and diagnostic classifications, and they do so under risk. This development process mirrors Bayesian reasoning through which teachers continually update their beliefs about effective teaching based on new evidence and insights, thus honing their predictive capabilities.

Shalem's framework (2014) illustrates how professional judgement develops through processes analogous to Bayesian reasoning. Teachers move beyond procedural knowledge to achieve the deeper professional judgement characterised by PCK2, enabling them to make informed predictions and decisions despite systemic uncertainty and cognitive constraints. This progression is essential for supporting genuine epistemological access for learners since teachers refine their strategies through ongoing belief updating and adaptation.

Understanding subjective probability in teaching

Bayesian probability offers a structured way to think about how prior knowledge is combined with new evidence to inform judgements under uncertainty. Rather than defining probability in terms of repeated trials or long-run frequencies, as in classical (frequentist) approaches (see Fisher, 1925), Bayesian reasoning treats probability as a measure of plausibility—how strongly one should believe a claim given the information currently available (Bayes, 1763). This makes it especially applicable to teaching, where the events of interest (such as whether a learner has understood a concept or will respond to an intervention) often cannot be observed repeatedly or measured directly. In such contexts, teachers must make forward-looking judgements by drawing on prior classroom experience, theoretical understanding, and immediate cues from learners—an approach well aligned with Bayesian reasoning (Shafto & Goodman, 2008). Rather than offering certainty, this perspective legitimates professional judgement as a rational, evidence-sensitive process, even when operating under conditions of ambiguity or limited feedback. The Bayesian framework provides three key elements that help explain professional judgement (Gleason & Harris, 2019):

- Prior Beliefs—$P(H)$: The initial subjective probability P of a hypothesis H based on existing knowledge;
- Likelihood—$P(E|H)$: The subjective probability of observing the evidence E given that the hypothesis H is true (how well the hypothesis predicts the evidence);
- Posterior Beliefs—$P(H|E)$: The updated subjective probability P of the hypothesis H after considering the new evidence E.

This structure mirrors how teachers develop professional judgement: they begin with theoretical knowledge and initial beliefs (priors), gather evidence through classroom experience (evidence), and update their understanding (posteriors) based on what they observe, essentially refining their internal predictive model.

While humans may not engage in precise mathematical computations of probabilities, the Bayesian framework can still model effectively how individuals intuitively update their beliefs in response to new information that challenges their predictions. Extensive research on cognitive biases and heuristics has demonstrated that human probability judgements often deviate from normative mathematical models (Kahneman et al., 1982). However, these intuitive judgements, although suboptimal from a purely mathematical perspective, are highly effective in real-world decisions where uncertainty and complexity are the norm (Gigerenzer et al., 2011).

This ecological rationality suggests that teachers use simple heuristics—mental shortcuts—to make quick decisions under pressure. These heuristics can be viewed as practical approximations of Bayesian updating, allowing teachers to adjust their beliefs and actions to reduce predictive error without complex calculations. Such fast and frugal rules of thumb, while not mathematically precise, are well-adapted to the specific challenges of classroom decision-making. Future research could explore specific heuristics teachers use (e.g., recognition, satisficing), but here I focus on how the underlying Bayesian principle models change in beliefs.

I position subjective Bayesian inference between simple heuristics and full-scale probabilistic models. It involves making intuitive judgements about what *might* work—drawing on heuristics—while systematically updating these judgements based on experience in a manner consistent with Bayesian principles (Parpart et al., 2018). This approach enables teachers to refine their professional judgement and predictive accuracy without relying on complex mathematical calculations.

Furthermore, the Bayesian framework is valuable for two key reasons: first, it provides a formal mechanism for understanding how feedback loops drive the probabilistic improvement of our beliefs. Subjective Bayesian reasoning outlines formally how these loops operate in teaching in probabilistic ways. When teachers implement a strategy (based on prior beliefs), they observe the outcomes (evidence). These observations act as feedback that informs whether their initial beliefs were accurate. By comparing predicted outcomes with actual results, teachers update their beliefs (posterior beliefs), refining their understanding of what strategies are effective (Tenenbaum et al., 2008). This updating mechanism offers insights into how professional judgement develops and improves over time, even when operating through simple intuitions rather than formal probabilistic reasoning. The point is that teachers do this in ways that work with best guesses that will probably work in the given conditions, and do these many times a day. We need a language that formally catches this, and Bayes assists with a first take.

Second, the Bayesian framework is useful because it helps explain both the stability and flexibility of professional judgement as characteristics of a well-calibrated predictive model. It shows how teachers can maintain coherent beliefs while continually updating them based on new evidence, integrating theoretical knowledge with practical experience systematically. This approach aligns with Luhmann and Schorr's (2000) emphasis on the probabilistic nature of teaching—acknowledging the inherent uncertainties—Simon's (1957) insights about bounded rationality—highlighting cognitive limitations and the use of heuristics—and Shalem's (2014) description of the development of sophisticated professional judgement through epistemological labour. By providing a formal mechanism for understanding how teachers develop and refine their professional judgement, the Bayesian framework enriches our comprehension of the complexities involved in teaching under uncertainty.

Bayesian models of teacher development: Using vignettes as a method

The application of Bayesian reasoning to teaching is risky for a broad readership in education since it involves probabilistic terminology that is off-putting to many. I use vignettes as constructed scenarios that take an imaginary narrative walk through the probabilistic terms to exemplify how the dynamics of probabilistic judgement unfold in practice (Klotz et al., 2021). Vignettes are "short, carefully constructed description of a person, object or situation, representing a systematic combination of characteristics" (Atzmüller & Steiner, 2010, p. 129). I present two constructed scenarios to illustrate how teachers develop professional judgement through intuitive Bayesian updating in response to specific classroom challenges: first, a classroom management scenario showing immediate belief updating done in an intuitive way; and second, a longer-term analysis with formal Bayesian terms of how teachers adopt and adapt scripted lesson plans (SLPs).

Constructed scenario 1: Classroom management decisions

Setting the Scene

It's third period on a Friday morning. Sarah, a new teacher in her first term, has been struggling with classroom management in her Grade 8 Science class. The staffroom has become her sanctuary during breaks, where experienced teachers offer advice, most of it centred on "establishing authority early" and "showing them who's boss." Today, after two previous challenging lessons, she's finally gotten her class settled for independent writing work. The room is quiet except for the scratch of pens, when suddenly, a rhythmic tapping breaks the silence. Lexi, a student in the middle row, is drumming her pencil.

Prior beliefs and multiple hypotheses

As Sarah confronts this moment of decision, her mind holds competing ideas about how to handle minor disruptions. These competing hypotheses—strict verbal intervention versus proximity control—represent what Bayesian theorists would call a probability distribution over possible actions and their predicted effectiveness. Note that in this account I am not providing hundreds of possible actions with strict probability distributions; it is a more natural process with the most obvious options present.

In the staffroom, her older colleagues have been clear: "These kids today need firm boundaries." Their voices echo in her mind, carrying the weight of institutional knowledge and experience. This advice has strongly biased her toward strict intervention, reflected in a high prior $P(H\_strict)$. Her own recent struggles with classroom management reinforce this belief: just yesterday, a gentle reminder to a chatting student seemed to undermine her authority further, leading to more disruption.

Yet somewhere in her consciousness, her university training persists. She remembers learning about proximity control, about subtle interventions that maintain student dignity and classroom flow. But these theories feel distant now, weighted down by daily reality and her growing need to establish herself as what she thinks of as a proper teacher. The chances of her using these gentler approaches, reflected in her prior $P(H\_gentle)$, have diminished with each challenging day.

Likelihood assessment

In the moment of Lexi's tapping, Sarah conducts what Simon would recognise as a bounded rationality analysis—a quick, intuitive assessment of likely outcomes under intense time pressure and cognitive constraints. She sees that the whole class can hear the tapping; she predicts a public response would likely demonstrate her control. The room's current quiet feels precarious and hard-won after previous struggles. She predicts a firm intervention could probably reinforce her authority, show students she will not tolerate disruption. Other students might be getting annoyed; they'd appreciate her taking action. It's an opportunity she predicts might set an example.

The possibility of proximity control flickers through her mind, but her likelihood estimates for its success—*P(Success | H_gentle)*—have been eroded by staffroom conversations and her own fears. It feels too risky, and the predicted outcome seems weak, might seem weak, might undermine the quiet she's finally achieved. The consensus among experienced teachers weighs heavily in her probability calculations, though she would not express it in these terms. Notice that the description of likelihood is forward looking; it is a projection of what is predicted to happen, given each hypothesis.

Evidence collection

Following her weighted probabilities (her current predictions), Sarah chooses the strict intervention: "Lexi, stop that tapping right now! We're trying to work!"

The evidence arrives with immediate clarity or what Bayesian theorists would call strong evidence creating a significant prediction error signal that challenges prior beliefs. Lexi jumps; she is startled and embarrassed. Several students look up from their work, the spell of concentration broken. Whispers break out in the back corner, the focused atmosphere dissipating like smoke. Two students start a side conversation, and Lexi looks confused and hurt since she hadn't even realised she was tapping.

Updating beliefs (Posterior)

This cascade of evidence forces Sarah to update her simple probability distribution over effective responses, adjusting her predictions about the effectiveness of public corrections. Her new understanding emerges not as abstract theory but as lived experience: public corrections can disrupt the whole class's learning more than the original problem; unconscious behaviours might need different approaches than does deliberate disruption; the cost of maintaining authority through strict intervention can outweigh its benefits; student embarrassment can damage both the learning environment and the crucial teacher-student relationship.

The next iteration

When similar tapping starts in the next lesson, Sarah's prior beliefs have shifted significantly given the updates (her predictive model has been adjusted). Though she still holds the possibility of strict intervention in her repertoire, her updated probability distribution leads her to predict proximity control might work better. She quietly moves near Lexi, gently touches her desk. The tapping stops, the class continues working undisturbed, Lexi remains engaged with her work, and the learning atmosphere is preserved.

This new evidence further refines her probability estimates and future predictions: minor, unconscious disruptions can be handled without sacrificing authority; maintaining learning momentum often matters more than public displays of control; different types of disruption require different responses; authority can be maintained through subtle interventions.

Through this process, Sarah begins developing professional judgement about classroom management. She learns that authority is not simply about strict control but about making strategic choices that balance immediate behaviour management with broader learning goals based on increasingly accurate predictions. She manages to distance herself from initial assumptions and then rework her assumptions to fit her experiences in a better way. Each new situation provides opportunities to refine these judgements further, leading to increasingly nuanced understanding of when to use different approaches.

This case illustrates how teachers, while not explicitly calculating probabilities, engage in intuitive Bayesian reasoning that allows them to refine their predictive judgement through experience. The process maintains coherent belief systems while incorporating new evidence, develops intuitive probability distributions over possible actions, and enables rapid yet nuanced decision-making in the dynamic classroom environment. It is a simple insight into how teachers weigh options when making decisions; there is no guarantee that any one method will work, and the number of methods are kept to a minimum. A few possible responses are weighed based on their probable success in the moment. Sarah's journey demonstrates how professional judgement evolves through systematic learning from experience, guided by an implicit Bayesian framework of prior beliefs, evidence collection, and belief updating driven by the need to improve predictions and outcomes.

Bridge between scenarios

The first case study, while illustrating the basic mechanics of Bayesian belief updating, operates primarily at the level of classroom management decisions. It shows how a novice teacher updates relatively simple beliefs about behavioural interventions but does not yet engage with the deeper aspects of professional judgement that constitute expert teaching. The case remains largely at the level of PCK1—the basic routines and management strategies that help teachers maintain classroom order and deliver lessons.

The following constructed scenario attempts to trace a more complex trajectory on how teachers, often reacting to the predictive failures of their initial approaches, might develop from PCK1 to PCK2 through their engagement with scripted lesson plans. By following a teacher's journey from university training through early classroom experiences with SLPs and beyond, I examine how Bayesian reasoning operates when teachers confront the fundamental challenges of predicting and enabling student learning rather than just managing behaviour. The constructed scenario is based loosely on De Clercq and Shalem's discussion of SLPs (2014), Shalem's (2017) work, Bertram's account of novice teachers developing experiences (2023), as well as the work of Sawyer (2001). In this account I will introduce Bayesian terms more formally.

Constructed scenario 2: The evolution of teacher beliefs: A Bayesian analysis of professional development

Intuitive account: Initial priors—The idealistic graduate

When Sarah graduated from her education program, her prior probability distribution reflected two distinct belief sets (initial predictions) about teaching effectiveness. Her university-acquired constructivist priors carried strong positive weightings; she considered student-centred learning highly likely to succeed along with problem-based instruction and creative exploration—*high P(H_constructivist).*

In contrast, her experiential priors from her own schooling carried much lower likelihood estimates, given her memories of how boring school was. Her prior probability distribution assigned relatively low likelihoods to the effectiveness of direct instruction and teacher-led discussions—*low P(H_traditional).* These traditional approaches, which she had experienced as a learner, were considered unlikely to produce optimal learning outcomes according to her university-influenced belief system (even though her university used lectures mainly as the delivery device).

Formal Bayesian analysis—Prior probabilities $P(H_1)$ and $P(H_2)$

In Bayesian terms, Sarah's initial beliefs (predictions) are represented by the prior probabilities $P(H_1)$ and $P(H_2)$, where:

- H_1: Hypothesis that constructivist approaches are effective.
- H_2: Hypothesis that traditional approaches are effective.
  Her prior $P(H_1)$ is high due to her university training, while $P(H_2)$ is low based on her personal experiences as a learner. The likelihood functions $P(E|H_1)$ and $P(E|H_2)$ represent the predicted probability of observing successful learning outcomes E given each teaching approach. Initially, $P(E|H_1)$ is considered high (she predicts success is likely with H_1), and $P(E|H_2)$ is considered low (she predicts success is unlikely with H_2). This reflects her strong confidence in constructivist methods over traditional ones as she enters the classroom.

Intuitive account: Confronting classroom reality: when theory meets reality

Sarah's confidence in student-centred teaching methods was quickly tested in the reality shock of early classroom experiences. A particularly memorable moment came during what she thought would be an engaging inquiry-based science lesson. She had planned carefully, creating open-ended exploration activities that matched perfectly her university training. However, the lesson descended into chaos: students were confused, off-task, and learning objectives were not met. Similar experiences followed with other constructivist lessons she attempted. Each unsuccessful attempt represented a significant prediction error, chipping away at her certainty about these methods, while more structured approaches started showing better results than she had expected. The stark contrast between her theoretical expectations and classroom reality forced her to reconsider rapidly her initial beliefs about what would work with real students.

Formal Bayesian analysis—Updating with likelihood $P(E|H)$ and posterior probability $P(H|E)$

Sarah's belief-updating process demonstrates Bayesian reasoning driven by prediction error:

- Likelihood Functions:
  - $P(E|H_1)$: The predicted probability of observing the classroom evidence E given constructivist methods H_1.
  - $P(E|H_2)$: The predicted probability of observing E given traditional methods H_2.

As she gathers evidence that constructivist methods are less effective than anticipated (the evidence makes $P(E|H_1)$ seem lower than predicted) and traditional methods are more effective (evidence makes $P(E|H_2)$ seem higher than predicted); she updates her beliefs.

- Posterior Probabilities:
  - $P(H_1|E)$: The updated probability that constructivist methods are effective given the evidence.
  - $P(H_2|E)$: The updated probability that traditional methods are effective given the evidence.

The likelihood ratio $P(E|H_2) / P(E|H_1)$ shifts dramatically with each new piece of evidence, leading to rapid updates in her posterior probabilities. Her initial priors lacked strong empirical grounding, so the new evidence (the prediction errors) significantly reshapes her beliefs about teaching effectiveness.

Intuitive account: Seeking stability—Turning to scripted lesson plans

After her early struggles (driven by predictive failures and resulting uncertainty), Sarah found comfort in the structured approach of scripted lesson plans. These detailed guides promised to solve her immediate challenges by offering predictability and providing clear, step-by-step instructions for every lesson over an extended period. For Sarah, these scripts offered a lifeline—a way to manage the overwhelming complexity of classroom life by providing clear structures and routines (strengthening her PCK1). They also began to offer insights into how to present content more effectively, subtly contributing to her ability to facilitate student understanding (initiating PCK2 development). The scripts reduced her cognitive load, freeing her from the need for constant difficult prediction and decision-making, letting her focus on delivery rather than constantly making decisions about content and pacing. SLPs gave her extensive scaffolding with PCK1 by providing pre-set lesson structures, explicit routines of work, along with assessment activities and worked out answers. Their apparent success with other teachers and their basis in systematic planning made them seem like a reliable solution to her current uncertainty about what would work in the classroom.

Formal Bayesian analysis—New prior $P(H_3)$, likelihood $P(E|H_3)$, and posterior $P(H_3|E)$

Sarah's adoption of SLPs represents a shift in her prior beliefs:

- New Hypothesis:
  - H_3: Hypothesis that Scripted Lesson Plans are effective.

Her new prior probability $P(H_3)$ is high due to the reputed success of SLPs and their structured nature. The likelihood function $P(E|H_3)$ represents the probability of observing successful outcomes E given the use of SLPs.

The posterior probability $P(H_3|E)$ reflects her strong belief that standardised procedures can increase teaching effectiveness, temporarily stabilising her shifting belief distribution by providing a seemingly reliable predictive model.

Intuitive account: Scripts meet students

As Sarah used the scripted lessons in her daily teaching, she began noticing important patterns (gathering new evidence). While the scripts worked reasonably well for introducing basic concepts in straightforward lessons, they often fell short when she was faced with the real complexities of her classroom. Her predictions based solely on the script were often inaccurate. Some students raced ahead of the prescribed pace while others needed more time. Key teachable moments were lost when she stuck rigidly to the script. Sometimes, the scripted explanations confused her students more than they helped. She found herself naturally adapting the scripts—actively testing modifications—adding extra examples for struggling students, skipping redundant steps for quick learners, and weaving in personal connections that engaged her specific class. This adaptation signified a deeper engagement with the content (developing her PCK2) while still relying on the structured framework provided by the scripts (utilising her PCK1). These daily experiences (new evidence challenging the simple model) showed her that effective teaching required more flexibility (a more nuanced predictive model) than pure script-following could provide. In Shalem's terms (2014), we see the emergence of a realised need for PCK2, the need for a more flexible and responsive understanding of what teachers need to do to enable student learning. Ironically, it is partly the cognitive space enabled by SLPs reducing cognitive load (less need for constant low-level prediction) that allows Sarah to engage with PCK2 (refining higher-level predictions about learning).

Formal Bayesian analysis—Revising likelihood functions $P(E|H_3)$ with conditional dependencies

Sarah's evidence collection led to significant revisions in her likelihood estimations for SLP effectiveness (her predictive model became more complex),

- Updated Likelihood Functions:
  - *$P(E|H_3$, student_type)*: Probability of success predicted given SLPs and different student types.
  - *$P(E|H_3$, content_difficulty)*: Probability predicted given SLPs and varying content difficulty.
  - *$P(E|H_3$, class_dynamics)*: Probability predicted given SLPs and specific class dynamics.

These conditional dependencies revealed that while strict script adherence (supporting her PCK1) provided lesson structure, it often resulted in lower likelihoods of deep student understanding (highlighting the need for further PCK2 development) than initially estimated (since the simple model failed to predict accurately across contexts). Her posterior probability $P(H_3|E)$ shifted toward a more nuanced view, where effectiveness probabilities depended heavily on contextual factors rather than just script fidelity.

Intuitive account: Developing professional insight—Beyond scripts

Sarah's journey led her to both discover and appreciate educative curriculum materials, resources that explained the *why* behind teaching decisions rather than just prescribing the *what* and the *how* (Shalem, 2017). Unlike rigid scripts, these materials helped her understand the reasoning behind different teaching approaches, enhancing both her lesson structuring skills (PCK1) and her ability to facilitate deep understanding (PCK2), effectively improving her predictive model. She found herself developing a deeper professional insight in knowing how to organise her lessons effectively and enable students to comprehend complex concepts, allowing her to adapt guidelines based on her students' needs (making better context-sensitive predictions). Rather than experiencing dramatic swings between different teaching approaches, she began making smaller, more refined adjustments to her practice in terms of fine-tuning her predictions and actions. When something did not work perfectly, she no longer abandoned it entirely (as she had with her early constructivist attempts) but, instead, thought carefully about how to modify it. Her understanding grew more stable but more nuanced: she could incorporate new experiences without completely overturning her existing knowledge (her predictive model became robust). In Morrow's (1994) epistemological access terms, a good teacher has in the back of their mind an understanding of what is involved in teaching (a sophisticated predictive model).

Formal Bayesian analysis—Joint likelihood function $P(E|H_4)$ and emergence of robust priors

Her evolving probability model demonstrates:

- New Hypothesis:
  - H_4: Hypothesis that combines theoretical understanding with practical adaptation (a more complex predictive model).
- Joint Likelihood Function:
  - $P(E|H_4)$: Predicted probability of success given both understanding and adaptation.

This refined model produced higher likelihood ratios than relying solely on SLP adherence. Her belief-updating process showed the emergence of robust priors—probability distributions that maintain their fundamental structure while accommodating new evidence through parameter updates rather than wholesale revision. Her predictions became more stable yet adaptable. Her new posterior probabilities $P(H_4|E)$ exhibited the stability characteristic of mature professional judgement.

Intuitive account: A measured approach to new teaching trends

When learning styles theory swept through her school, Sarah responded differently from how she would have as a new teacher. Instead of immediately embracing this new approach with enthusiasm or rejecting it outright, she took a measured approach. Her hard-won experience and daily epistemological labour had taught her to evaluate carefully new ideas against her existing predictive model. While colleagues rushed to categorise students as visual, auditory, or kinaesthetic learners, Sarah tested these ideas cautiously in her classroom, comparing results with her existing effective practices. She found herself asking, "How does this fit with what I already predict works?" rather than seeing it as a complete solution. She did not go and read the latest research and empirical evidence on learning styles theory—she did not have the time or inclination for this—but took what was useful pragmatically if it improved her predictions/outcomes without radically changing her practices.

Formal Bayesian analysis—Incorporating sceptical priors P(H) and evaluating new evidence with likelihood ratios

Sarah's matured Bayesian framework involved:

- Sceptical Prior:
    - A prior probability $P(H)$ that requires robust evidence to significantly shift her beliefs (her predictive model is stable).
- Evaluating New Hypotheses:
    - H_new: Hypothesis representing the new teaching trend (e.g., learning styles theory).
    - Likelihood Functions:
        - $P(E|H\_new)$: Predicted probability of observing evidence given the new approach.
        - $P(E|H\_proven)$: Predicted probability given proven methods.
- Likelihood Ratio:
    - $P(E|H\_new) / P(E|H\_proven)$: Used to compare the predicted effectiveness of new approaches against established ones.

Her strong professional judgement enabled gradual belief updating rather than dramatic shifts, maintaining high posterior probabilities $P(H\_proven|E)$ for effective practices while methodically evaluating new claims.

Intuitive account: Achieving confident flexibility—The seasoned professional

After years of teaching, Sarah developed a deep, stable sense of what works in her classroom. Her understanding was not rigid; it was flexibly responsive to new situations while remaining grounded in proven practices (a robust and refined predictive model). She could assess quickly whether a new approach would be likely to succeed with particular students or topics, drawing on her rich experience without being trapped by it. When faced with novel teaching challenges, she neither clung desperately to old methods nor jumped impulsively to new ones. Instead, she adapted her approach thoughtfully based on the specific context in making nuanced predictions about the students involved, the subject matter at hand, and the particular learning goals. When reading research on learning and *how to* books on teaching, she looked for ways to incorporate these insights into her teaching practices, but almost never in a wholesale way and always integrating them to refine her predictive model). Her professional judgement had become more reliable and more nuanced, reflecting the mature integration of PCK1 and PCK2. This allowed her to navigate new situations with confident flexibility—structuring her lessons effectively (PCK1) while understanding deeply how to facilitate student learning of complex material (PCK2)—all while maintaining what she knew worked (relying on her well-calibrated predictive model).

Formal Bayesian analysis—Mature robust priors P(H) and evolved likelihood function P(E|H)

Sarah's mature professional judgement is reflected in:

- Robust and Refined Priors:
    - *P(H)* distributions that maintain structural stability while incorporating new evidence (a stable yet adaptable predictive model).
- Evolved Likelihood Function:
    - *P(E|H)* that handles complex combinations (allowing for nuanced predictions):
        - *P(E|H) = P(success | strategy, context, student_needs, subject_matter)*

Her probability framework generated high posterior probabilities for approaches balancing structure with adaptability. The key feature of her evolved belief system was its ability to produce accurate probability estimates for novel situations without dramatic distribution shifts based on limited evidence. Her likelihood function *P(E|H)* demonstrated professional maturity by processing new information effectively (refining predictions) while maintaining responsiveness to substantive evidence of effective approaches.

Conclusion

I have argued that Bayesian reasoning offers a probabilistic framework for understanding how professional judgement develops as teachers learn to navigate uncertainty through prediction and belief updating. Probabilistic thinking helps teachers navigate the inherent uncertainties of education while working within cognitive constraints. Subjective Bayesian analysis formalises the mechanisms of belief updating that underlie heuristic decision-making processes by modelling how prior beliefs are adjusted in response to prediction errors observed in practice.

Luhmann and Schorr's (2000) insight about the separation between teaching and learning systems highlights why teachers must think probabilistically; they can influence but never directly control student learning, requiring them to constantly update predictions about effectiveness. Simon's (1957) bounded rationality framework explains why teachers must develop efficient predictive heuristics rather than seek optimal solutions, while Shalem's work (2014) shows how these heuristics can become increasingly sophisticated as teachers build the knowledge structures underpinning effective prediction through systematic development of professional knowledge. While this development may not always align with formal academic classifications, it represents a sophisticated form of professional judgement adapted to the practical demands and constraints of classroom teaching.

The Bayesian framework helps explain how teachers progress from being novices to experts. Novice teachers often start with fragile priors—initial beliefs based primarily on theoretical knowledge and personal learning experiences, which may lack robust empirical grounding. This leads to easily disrupted predictions and dramatic belief updates when they are confronted with classroom realities (significant prediction errors), explaining phenomena like the rapid abandonment of learner-centred methods in favour of more structured approaches. As teachers gain experience, they develop more robust and refined priors—stable predictive models that can be updated efficiently with new evidence without requiring drastic changes to their overall framework. This development allows teachers to refine their professional judgement and predictive accuracy efficiently while maintaining a consistent and adaptable teaching approach.

Postscript

A key motivation for this analysis stems from my interest in developing a more general theory of pedagogy—one that can account for teaching to learn processes in both humans and, increasingly, artificial intelligence systems. It concerns the fundamental question of how we teach to facilitate learning in both biological and computational systems. Central to this project is the subjective Bayesian reasoning explored in this paper. It provides a powerful, abstract language for describing how systems manage uncertainty, make predictions, and update their internal models based on experience. This probabilistic perspective applies remarkably well to modelling cognitive processes in humans, as argued here for teacher development, and it is also foundational to how many modern AI systems learn. I chose not to weigh down the main body of this paper with such a detailed juxtaposition since my primary aim was to establish the utility of the Bayesian framework within the familiar context of teacher professional development. I know that Bayesian reasoning is not what most readers of the *Journal of Education* are familiar with but really do feel that it is a vital conceptual tool for us to learn as AI progressively becomes a part of our lives. It cannot be that our only major question is how we work with AI in teaching and researching. We need to ask about the fundamental processes behind how AI is taught to learn and what this means for a more general theory of pedagogy.

AI systems, particularly those designed for learning and adaptation, can be understood as engaging in ongoing probabilistic inference. They construct and maintain internal models that represent their so-called beliefs about the world—often structured as probability distributions (e.g., the probability that an image contains a certain object, the probability of the next word in a sentence). Learning, in this context, is the process of refining these internal probabilistic models based on interaction with data or an environment.

This learning process fundamentally aligns with Bayesian principles. An AI system starts with an initial model, reflecting its prior beliefs—$P(H)$. These priors can originate from various sources—the vast amounts of data used in "pretraining," explicit rules programmed by designers, or even an initial state of relative ignorance (e.g., uniform probability distributions). As the AI encounters new data or receives feedback—collectively termed evidence E—it must integrate this information. The critical step involves evaluating how likely this new evidence was, given the system's current model. This is captured by the likelihood—$P(E|H)$. If the evidence strongly contradicts the model's predictions, it signifies a large "prediction error" or high "surprise."

Bayes' Theorem provides the formal mechanism for updating the system's beliefs. It mathematically combines the prior belief—$P(H)$—with the likelihood of the new evidence—$P(E|H)$—to compute an updated belief, the posterior probability—$(P(H|E)$. This posterior represents a revised internal model that better accounts for the recent evidence. Through repeated cycles of encountering evidence and updating beliefs, the AI's model becomes progressively more accurate in its representation of the world and its ability to make predictions. Many sophisticated AI algorithms, including those in deep learning, implement efficient methods to approximate this Bayesian updating cycle, enabling learning even in highly complex scenarios (Gal & Ghahramani, 2016).

Reinforcement learning (RL) in AI offers a clear example of this in action. An RL agent learns to make sequences of decisions in an environment to maximise cumulative rewards. Its current strategy for choosing actions in different situations can be seen as its policy, reflecting its prior beliefs about optimal behaviour. When it takes an action and observes the outcome (a new state and a reward signal), this serves as evidence. The agent calculates a prediction error—the difference between the expected reward and the actual reward received (related to the likelihood). This error signal drives adjustments to its policy, using algorithms functionally equivalent to Bayesian updating, leading to a posterior policy that incorporates the new experience. This iterative process allows the agent to learn effective strategies through trial-and-error, guided by probabilistic updates informed by feedback (Sutton & Barto, 2018). This mirrors the teacher development process, where practical experience and feedback drive updates to pedagogical beliefs and strategies, although there are dramatic differences as well.

The convergence of Bayesian principles in modelling learning across diverse fields from the AI techniques discussed here (Pearl, 1988; MacKay, 2003) to cognitive neuroscience, where the brain is conceptualised as a "Bayesian prediction machine" actively working to minimise prediction error (Friston, 2010, 2013) suggests strongly that we are identifying fundamental, abstract mechanisms of intelligence and adaptation. For educational theory, this is significant. It opens the door to developing a more unified science of learning and instruction. By focusing on these underlying principles—how systems represent uncertainty, make predictions, and update beliefs based on error—potentially we can identify pedagogical strategies that are effective precisely because they align with these core learning dynamics, irrespective of whether the learner is human or artificial. This pursuit echoes the ambitions of Cybernetics (Wiener, 1948; Pask, 1961) to discover general principles governing complex adaptive systems. Subjective Bayesian reasoning provides a robust and contemporary framework for continuing this important work, leading potentially to breakthroughs in both how we understand learning and how we design effective teaching for all types of learners in an increasingly integrated future.

References

Abbott, A. (1988). *The system of professions: An essay on the division of expert labour*. University of Chicago Press.

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *6*(3), 128–138. https://doi.org/10.1027/1614-2241/a000014

Baraldi, C., Corsi, G., & Esposito, E. (2021). *Unlocking Luhmann: A keyword introduction to systems theory* (K. Walker, Trans.). Bielefeld University Press. https://doi.org/10.14361/9783839456743

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society*, *53,* 370–418.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher*, *15*(7), 5–13.

Berliner, D. C. (1987). Ways of thinking about students and classrooms by more and less experienced teachers. In J. Calderhead (Ed.), *Exploring teachers' thinking* (pp. 60–83). Holt, Rinehart & Winston.

Bishop, A. J., & Whitfield, R. C. (1972*). Situations in teaching*. McGraw-Hill.

Bertram, C. (2023). A systems approach to understanding novice teachers' experiences and professional learning. *Journal of Education, 90,* 11–31. https://dx.doi.org/10.17159/2520-9868/i90a01

Bobadilla-Suarez, S., Jones, M., & Love, B. C. (2022). Robust priors for regularized regression. *Cognitive Psychology*, *132*, 101444. https://doi.org/10.1016/j.cogpsych.2021.101444

Borko, H., Roberts, S., & Shavelson, R. J. (2008). Teachers' decision making: From Alan J. Bishop to today. In P. Clarkson & N. Presmeg (Eds.), *Critical issues in mathematics education* (pp. 37–67). Springer. https://doi.org/10.1007/978-0-387-09673-5_4

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81.

Deacon, R. (2016). The initial teacher education research project: Final report. https://www.jet.org.za/clearinghouse/projects/iterp/resources.

De Clercq, F., & Shalem, Y. (2014). Teacher knowledge and employer-driven professional development: A critical analysis of the Gauteng Department of Education programmes. *Southern African Review of Education*, *20*(1), 129–147.

Fisher, R. A. (1925). *Statistical methods for research workers*. Oliver and Boyd.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. (2013). Active inference and free energy. *Behavioral and Brain Sciences*, *36*(3), 212–213. https://doi.org/10.1017/S0140525X12002142

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In M. F. Balcan & K. Q. Weinberger (Eds.), *Proceedings of the 33rd International Conference on Machine Learning* (Vol. 48, pp. 1050–1059). PMLR. https://doi.org/10.48550/arXiv.1506.02142

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, *521*(7553), 452–459. https://doi.org/10.1038/nature14541

Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford University Press.

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.

Gleason, P., & Harris, J. (2019). The Bayesian approach to decision making and analysis in nutrition research and practice. *Journal of the Academy of Nutrition and Dietetics*, *119*(10), 1676–1680. https://doi.org/10.1016/j.jand.2019.07.009

Gourlay, A., Mshana, G., Birdthistle, I., Bulugu, G., Zaba, B., & Urassa, M. (2014). Using vignettes in qualitative research to explore barriers and facilitating factors to the uptake of prevention of mother-to-child transmission services in rural Tanzania: A critical analysis. *BMC Medical Research Methodology*, *14*(21). https://doi.org/10.1186/1471-2288-14-21

Joyce, J. M. (2004). The development of subjective Bayesianism. In D. M. Gabbay, J. Woods, & A. Kanamori (Eds.), *Handbook of the history of logic* (pp. 10–415). Elsevier.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Klotz, S., Kratzer, S., Westner, M., & Strahringer, S. (2021). Literary sketches in Information Systems research: Conceptualization and guidance for using vignettes as a narrative form. *Information Systems Management*, *39*(4), 345–362. https://doi.org/10.1080/10580530.2021.1996661

Lee, O., & Porter, A. C. (1990). Bounded rationality in classroom teaching. *Educational Psychologist*, *25*(2), 159–171. https://doi.org/10.1207/s15326985ep2502_4

Luhmann, N. (1986). "The Autopoiesis of social systems." In F. Geyer & J. Van d. Zeuwen (Eds.), *Sociocybernetic Paradoxes: Observation, Control and Evolution of Self-Steering Systems* (pp. 172–192). SAGE.

Luhmann, N. (1992). Operational closure and structural coupling: The differentiation of the legal system. *Cardozo Law Review*, *13,* 1419–1441.

Luhmann, N. (1995*). Social systems*. (J. Bednarz Jr. & D. Baecker, Trans.). Stanford University Press. (Original work published 1984)

Luhmann, N. (2012). *Introduction to systems theory*. (P. Gilgen, Trans.). Polity Press. (Original work published 2002)

Luhmann, N., & Schorr, K. E. (2000). *Problems of reflection in the system of education*. Waxmann.

MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press.

Morrow, W. (1994). Entitlement and achievement in education. *Studies in Philosophy and Education*, *13*(1), 33–47. https://doi.org/10.1007/BF01074084

Parpart, P., Jones, M., & Love, B. C. (2018). Heuristics as Bayesian inference under extreme priors. *Cognitive Psychology, 102*, 127–144. https://doi.org/10.1016/j.cogpsych.2017.11.006

Pask, G. (1961). *An Approach to Cybernetics*. Hutchinson.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers.

Qvortrup, L. (2024). The impossibility and necessity of causality in Niklas Luhmann's theory of education. *Educational Theory*, *73*(6), 917–937. https://doi.org/10.1111/edth.12608

Sawyer, R. D. (2001). Teacher decision-making as a fulcrum for teacher development: Exploring structures of growth. *Teacher Development*, *5*(1), 39–58. https://doi.org/10.1080/13664539700200133

Shafto, P., & Goodman, N. D. (2008). A Bayesian model of pedagogical reasoning. In *Proceedings of the 2008 AAAI Fall Symposium on Naturally-Inspired Artificial Intelligence* (Technical Report FS-08-06, pp. 101–102). AAAI Press.

Shalem, Y. (2014). What binds professional judgement? The case of teaching. In M. Young & J. Muller (Eds.), *Knowledge, expertise and the professions* (pp. 93–105). Routledge. https://doi.org/10.4324/9781315883083

Shalem, Y. (2017). Scripted lesson plans: What is visible and invisible in visible pedagogy? In B. Barrett, U. Hoadley & J. Morgan (Eds.), *Knowledge, curriculum and equity: Social realist perspectives* (pp. 183–199). Routledge. https://doi.org/10.4324/9781315111360

Shalem, Y., & Slonimsky, L. (2010). Seeing epistemic order: Construction and transmission of evaluative criteria. *British Journal of Sociology of Education*, *31*(6), 755–778.

Slonimsky, L., & Shalem, Y. (2006). Pedagogic responsiveness for academic depth. *Journal of Education*, *40*, 35–58.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Simon, H. A. (1957). *Models of man: Social and rational*. John Wiley and Sons.

Simon, H. A. (1976). From substantive to procedural rationality. In T. J. Kastelein, S. K. Kuipers, W. A. Nijenhuis, & G. R. Wagenaar (Eds.), *25 years of economic theory* (pp. 65–86). Springer. https://doi.org/10.1007/978-1-4613-4367-7_6

Simon, H. A. (1992). "What is an 'explanation' of behavior?" *Psychological Science*, *3*(3), 150–161.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285. https://doi.org/10.1126/science.1192788

Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine*. MIT Press.

Wilson, A. G., & Izmailov, P. (2020). Bayesian deep learning and a probabilistic perspective of generalization. *Advances in Neural Information Processing Systems*, *33*, 4697–4708. https://doi.org/10.48550/arXiv.2002.08791

Winch, C. (2017). *Teachers' know-how: A philosophical investigation*. Wiley-Blackwell. https://doi.org/10.1002/9781119355700