New Directions for Evaluation

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Reasoning in Evaluation: Why Does It Matter?

Benjamin Harris | Lyn Alderman

Academic Transformation Portfolio, Academic Division, University of Southern Queensland, Toowoomba, Queensland, Australia

Correspondence: Benjamin Harris (benjamin.harris@usq.edu.au)

Received: 16 April 2025 | **Accepted:** 17 April 2025

ABSTRACT

This article addresses two significant philosophical debates of importance to the evaluation context, which have transcended science and social sciences for centuries. These debates center around philosophical approaches and approaches to reasoning as they relate to evaluation. We aim to impress upon evaluators the importance of understanding how and why philosophical approaches; specifically holism and reductionism, and approaches to reasoning; specifically deductive, inductive, and abductive approaches, inform their general practice and approaches to individual evaluations. This article does not intend to weigh in on or add to these debates, as they have received sufficient treatment elsewhere by philosophers and logicians. Rather, we argue that within their situational context, evaluators must understand their philosophical choices and their approaches to reasoning and the impact that choices and approaches will have on the evaluation activities. Hence, the aim of this article is to explore situational context, philosophical choices, approaches to reasoning as a decision-tree that leads to sound, effective methodological choices and useful evaluation outcomes. Like all articles in this issue, we do not believe any of these approaches are better or more preferable to evaluation in general, rather, we encourage evaluators to understand the choices and approaches they have at their disposal and to leverage their choices in answering their evaluation questions.

1 | Introduction

Evaluation textbooks and journals are filled with a multitude of methodologies and methods that can be applied to evaluation. This is hardly surprising given that many leading figures in the field consider methodological expertise to be a key prerequisite to being a good evaluator (Lemire et al. 2020; Mertens 2024; Patton 2011; Patton and Campbell-Patton 2024; Scriven 1994; Weiss 1998). However, what is less frequently discussed is the choices evaluators need to make prior to methodology—that is their approach to reasoning (deductive or inductive) and whether to assess programs as a whole or by their individual components (the holism-reductionism debate). The discussion that follows will outline these two important debates and explain why they should be a central concern of every evaluator's practice.

In the Australian context (see Alderman and Harris 2025), we have observed weak signals that suggest the historical culture of holistic approaches and inductive reasoning may be eroding. With the exponential increase in data collection and storage over the past few decades, there is increased pressure on those designing an intervention to know what is going to happen before it happens (Bray et al. 2020), setting the scene for future evaluators with systematic theories of change and logic models. Moreover, with so much data now available and the rise of numerous methodologies that allow us to quantify traditionally unquantifiable metrics, evaluators have new opportunities to generate deeper insights. For example, with the rise of "big data" analytics evaluators are now able to identify trends and patterns in large textual datasets (Raveh et al. 2020). This tends to result in experiments or quasi-experiments taking place, lending themselves to a deductive approach to reasoning. While we do not advocate for one style of reasoning over the other, there is a sense that inductive approaches are becoming less popular (see Alderman and Harris's Editor's Notes in this special issue). As such, we are concerned this may lead to an overreliance on

deductive approaches to explain complex, multifaceted phenomena. This is at odds with some of evaluation's leading voices, such as Scriven (1991), who argued for goal-free evaluation, which Thomas (2006, 238) has since determined as being "consistent with an inductive approach." Stufflebeam (1968) and Patton (1978) both alluded to inadequate consideration of the commissioner's context and their intentions. Ultimately, the evaluation must resonate with and be meaningful to those who commissioned it, not just the evaluator. Hence, understanding the context of the evaluation is critical to evaluation decision-making processes.

2 | Revisiting Context

Before discussing the deductive-inductive reasoning and holism-reductionism debates, it will be instructive to revisit the idea of context-informed evaluation. As the rest of this article will detail, making choices about the philosophical approach and approach to reasoning without an understanding of the contexts in which the evaluand takes place is nonsensical. Context has become a buzzword commonly associated with evaluation and social science research. Yet despite this, there are barely any definitions, let alone agreed ones. In evaluation circles, context is often quickly conflated with Pawson and Tilley's (1997) realist evaluation. The now well-known context-mechanism-outcome chain provides evaluators with a systematic method to determine whether contextual changes bring about different mechanisms and/or outcomes. However, context in evaluation should be thought of more broadly and has a longer history than realist evaluation alone.

Greene (2005, 83) provides the foundational definition in the evaluation context, stating, "context refers to the setting within which the evaluand (the program, policy, or product being evaluated) and thus the evaluation are situated. Context is the site, location, environment, or milieu for a given evaluand. It is an enormously complex phenomenon." This definition was later followed by a long list of dimensions that make up context. Rog (2012, 26) argued a "context-first approach to evaluation is more appropriate" than a "methods-first approach," thus supporting the notion that understanding the context should be the first step in any evaluation. Rog (2012) expanded on Greene's ideas, organizing context into five domains—the problem, the intervention, the broader environment, decision-making, and evaluation. As such, Rog (2012) argued that understanding the context of these five domains leads evaluators to methods appropriate to the context of each individual evaluand. This line of thinking has since been endorsed and expanded upon by several evaluation researchers (Coldwell and Moore 2024; Horne 2017; Newton-Levinson et al. 2020; Harris et al. 2025). While there is obvious agreement over what context is and why it is important, only a few have provided guidance on how to uncover context (see Conner et al. 2012; Harris et al. 2025). This article suggests that an additional step should be taken by evaluators following contextual understanding and before methodological decisions, that is using context to help inform the most appropriate philosophical approach and approach to reasoning. As such, prior to methodological choices and evaluation design, evaluators should have a deep understanding of:

- The context in which the system, program, or sub-elements are situated.
- The philosophical approaches at play in terms of holism or reductionism.
- The advantages of disadvantages of selecting deductive, inductive, and abductive reasoning.

Therefore, when an evaluator is informed of the context, philosophical approaches, and reasoning, this will allow for the most appropriate selection of methodologies to build into the evaluation design. Commissioners of evaluations, where appropriate, should provide potential evaluators with this contextual information up front to inform sound and rigorous evaluation design.

Once context is established, it is important for evaluators to understand *what* they are evaluating and *how* they will draw their conclusions. Some of this discussion can be found in the already well documented evaluation literature on the concept of "evaluative thinking," including the 2018 issue of this journal titled *Evaluative Thinking* which touches on deductive-inductive reasoning debate and its role in evaluation (Buckley et al. 2015; Cole 2023; Patton 2018, 2019; Vo et al. 2018). However, evaluative thinking transcends and goes beyond the philosophical choices evaluators should be consciously making before starting. In the literature, evaluative thinking is applied throughout the entire way through evaluation process. It is distinct from other forms of thinking, and arguably more complex than standard approaches to reasoning and thinking, because it must draw conclusions based on judgements of value and utility (Cole 2023).

While the discussion within this article might sound exceedingly obvious, these questions of what and how have a deeper level. The *what* question is concerned with the holism-reductionism debate and speaks to the scope of the analysis that will take place during the evaluation. The *how* question is centered around the deductive-inductive reasoning debate, specifically the logic and approach to reasoning which dictates how conclusions are reached. As the rest of this article will detail, these are important considerations for the evaluator to be consciously aware of, as they will impact the nature of the results.

3 | The Holism-Reductionism Philosophical Debate

There is a broad consensus that Smuts (1926) coined the term holism in his book *Holism and Evolution*. While sometimes criticized for its lack of philosophical rigor, the basic argument that nature consists of more than the sum of its parts laid the foundation for holism going forward. Contemporarily, holism provided the foundation for modern systems theory, a belief that systems (whether they be biological, ecological, information, or cultural) are greater than their individual parts (Laszlo and Krippner 1998; Midgley 2006; Senge 1990, 1997). More specifically, systems are not just the assembly and connection of individual components; phenomena take place because of the system as a whole and would not take place if each component operated separately. Smuts (1926, 122) explains, "a whole is not the parts

which compose it; it is these parts in their intimate union and the new reactions and functions which result from that union."

The term reductionism does not have a clear origin story and has not been credited to a single individual. However, reductionist thinking has clearly been in existence for millennia as ancient philosophers sought to break down their observations into simple and understandable scientific laws to explain specific phenomena. From a systems perspective, Rene Descartes is often cited as a key figure in reductionist history, citing several examples from the natural world of systems merely being machines comprised of smaller parts (Hutchins 2015). Pure reductionism sits at the opposite end of the spectrum, arguing that everything (objects, events, and organizations) are made up of individual, often minute, elements, which cannot be broken down any further (Ackoff 1974). In the hard sciences, this would be akin to subatomic particles. In the world of humanities, this would mean individual human beings rather than tribes, nations, or other collections of people.

Furthermore, reductionists would argue that these elements each have their own function and therefore should be studied and considered in isolation from other elements. While reductionists acknowledge that the natural and physical worlds consist of systems, they reject the idea that these systems possess a mystical or unseen ability to transform their parts into an indivisible whole. Instead, they believe even closely linked systems can be explained by the interactions and operations of their individual components (Verschuren 2001). If not already clear, holism and reductionism in their purest forms are extremes that very rarely mirror the realistic approach that researchers and evaluators take in practice. In fact, the extremes would likely reflect poor evaluative practice. Not considering individual components of complex systems and interactions does not allow for results to show which factors of a program are more impactful than others. By contrast, everything in isolation will not consider the dependencies and interactions that alter the impacts of individual components. As such, it is more useful to consider holism and reductionism as two extremes of a spectrum where most approaches will be more centered, with tendencies toward holistic and reductionist paradigms. Put another way, Verschuren (2001) argues pure holism is near impossible to achieve and argues "less reductionist" is probably a more accurate means of describing a holistic approach to research and problem-solving.

The holism-reductionism debate is not distinctly discussed within evaluation literature to date. However, it can be detected in some of the broader approaches of certain individuals, particularly system approaches to evaluation. Patton's (2011) developmental evaluation does provide a critical undertone of reductionist approaches such as systematic logic models, which attempt to draw connections between highly specific activities and discrete short-term outcomes. Moreover, Patton acknowledges most programs are operating amidst multiple systems, all of which are complex and overlap. Therefore, trying to determine linear, causal relationships at highly discrete levels is likely to be flawed. House (1993) critiques purely quantitative and reductionist evaluation approaches, cautioning that these methods overlook stakeholder perspectives, diverse understandings of value, and the importance of contextual factors. It should be noted, however, that quantitative analysis should not always be equated with reductionism. Essentially, House argues that individuals interact with systems and their component parts differently. As such, an overly reductionist approach would not provide a realistic view of how most individuals experience phenomena. Even Albert Einstein, while never using the term "holism," has been described by his contemporaries as a believer in holism over reductionism (Caruana 2005; Renn and Gutfreund 2023). Correll et al. (2014) also believe there is a role for both holism and reductionism in evaluation, and like us, suggest context should drive the choice. A particular reading of Pawson and Tilley's (1997) realist evaluation paradigm can be viewed as a means of combining the best tenets of holism and reductionism. By forcing evaluators to consider how and if context changes outcomes when looking for causal relationships between interventions and outcomes, it considers the explanation for outcomes more holistically. However, one could also argue that segmenting out each specific context within an evaluation to deliver segmented results could be viewed as highly reductionist.

Again, the purpose of this discussion is not to direct evaluators toward holism or reductionism, but rather it is important to understand which approach is being taken and why. These decisions will often be guided by the discussion's evaluators have with commissioners. For example, some commissioners will have highly specific questions that the funders require answers to and evidence for to continue. In these instances, highly reductionist approaches that control for variables and segment out individual components of interventions would be ideal. By contrast, a commissioner looking to challenge the same funders to see value and impact beyond the participants may warrant a more holistic approach, which views the participants as an indivisible component of a larger system, such as a community. Furthermore, they will deliver different results when examining the same program. It is worth noting that there are a number of other philosophical decisions that could be posed, which are outside the scope of this article. However, if evaluators address the holistic/reductionism decision alone, this will enhance the evaluation design and selection of methodologies.

4 | The Deductive-Inductive Reasoning Debate

An additional choice evaluators need to make is related to reasoning. Reasoning is essentially the logic of making arguments and drawing conclusions. Logical and coherent reasoning is important for any task that involves reaching conclusions, whether it be research, investigation, or evaluation. As pointed out earlier, because evaluation is arguably more complex than research, given conclusions of value need to be reached, clarity on the reasoning that leads to it is important for the evaluator to understand and should ideally be clearly communicated to the evaluation's audience (Cole 2023). Reasoning is a complex subject, and multiple paradigms for reasoning exist, and books could fill the typology of all reasoning approaches. As such, this section will limit discussion to the most recognized forms—deductive, inductive, and abductive reasoning—and how these impact evaluations.

Both deductive and inductive reasoning have their place in evaluation practice. One is not preferable over the other. As detailed

above, every evaluation has its own context, and therefore, the decision should be guided by what is most appropriate for the context. Deductive and inductive reasoning are not dissimilar to holism and reductionism, in the sense that they are essentially opposites of one another. However, it does not operate on a spectrum. Reasoning cannot be "somewhat" deductive or inductive. Although it is possible to situate small deductions within broad inductive approaches and vice versa.

Deductive reasoning has its roots in the scientific method. Aristotle has largely been credited as the first to describe the deductive process. His famous example, "Socrates is a man; all men are mortal; therefore, Socrates is mortal" remains a simple and elegant illustration used in encyclopedias and textbooks to date (Shanahan 1989; Schechter 2013; Wilson 2020). As it shows, deductive reasoning starts with a premise (or hypothesis) which is based on a predetermined truth and therefore communicated as an absolute statement. The investigation and data collection are guided by searching for evidence that validates or invalidates the initial hypothesis (Creswell and Clark 2017).

The deduction/induction debate has existed for 400 years since Francis Bacon articulated the concept of induction in his 1620 work "New Organon." Inductive reasoning requires data collection and analysis guided by no preconceptions. The investigator simply starts to gather data and evidence on a particular topic with no notion of attempting to validate a preconception. The data analysis then moves toward generalizations that explain the observations (Klauer and Phye 2008). While not necessarily remembered for inductive reasoning, Wittgenstein's (1958) phrase, "don't think, but look!" has become a mantra for inductive reasoning. It is a simple and elegant mantra that encapsulates the spirit of observation, surpassing a preconceived notion or testing of a hypothesis.

Inductive reasoning is not constrained by the scope of a hypothesis—everything is relevant until deemed otherwise. This gives space to unexpected and unintended results to be uncovered and explored by the evaluator that would otherwise be missed with a deductive approach. Context is everything when it comes to evaluation, and context is often missed in a myopic, hypothesis-driven approach. "What" questions are typically the questions that drive evaluations and are also the basis of inductive reasoning, for example: "What is happening here?"; "What is the history and background of this program?"; "What is that person doing?" These are all standard questions that are either overtly asked by evaluators or form part of their reasoning throughout the evaluation. A more technical explanation of the distinction is that deduction is confirmatory and is associated with the positivist research epistemology. By contrast, induction is exploratory and connected to the interpretivist school of epistemology (Williamson 2002). Deductive reasoning provides a logical process through which we add to or strengthen existing knowledge by affirming or nullifying a hypothesis. This is one of the inherent strengths of deduction and largely why it has been preferred by academics for centuries (Thomas 2021). However, its greatest limitation is the reliance on the hypothesis being based on knowledge with a truth value (Thomas 2021). Should such knowledge be found to be lacking truth value at a time in the future, all associated reasoning based on it becomes invalidated. Inductive reasoning also has issues. The major problem is the inability to validate generalizations and conclusions, as they are not based on a predetermined truth. Moreover, there is an assumption that all future observations will look the same as past observations (Hume 2000). Therefore, from a logical perspective, generalizations can only ever be probabilistic.

Reasoning eventually results in arguments. Here, logicians provide further guidance on how these arguments need to be formed. Deductive arguments need to be both valid and sound. Gensler (2010, 2) provides a useful example to determine validity in that a valid argument might be, "if you oversleep, you'll be late. You aren't late, you did not oversleep." The premise of this argument suggests that oversleeping will result in lateness. Lateness did not occur; therefore, we can conclude that oversleeping did not take place.

Comparatively, the following argument is invalid: "if you oversleep, you'll be late. You didn't oversleep, you aren't late" (Gensler 2010, 3). On initial inspection, this does not appear problematic. However, the construction of the premise does not make oversleeping is the only variable for lateness. There are a number of other things that could take place resulting in lateness. Therefore, the argument is invalid. For arguments to be sound, the premise has to be true. When it comes to deductive arguments, there is an assumption that the premise is true. This assumption comes about from something predetermined to be factually accurate. Therefore, there can be, and historically often have been, instances where the argument is perfectly valid (at the time it is presented), but later the premise was proven to be false, making the argument not sound and not valid. The arguments are presumably informed by prior tests, which have validated or invalidated the hypothesis. For example, in Gensler's oversleeping and lateness argument, if a series of experiments is conducted to test the argument and each experiment results in lateness, the hypothesis is proven to be correct. This validation of the argument then goes on to act as more compelling evidence for future researchers to use the same hypothesis. However, if even one of the experiment's results is being on time, the hypothesis fails.

Inductive arguments mirror deductive arguments in terms of flow and progression. Rather than starting with a hypothesis and collecting data to prove (or disprove) that hypothesis, the arguments start with specific data points, which can be generalized into a hypothesis (Hayes and Heit 2018; Heit 2007). This data can essentially be anything, but the point is best demonstrated by observational data. For example, if I go to the park and only see white ducks swimming in the lake, I can conclude that most ducks are likely to be white. The difference is that the hypothesis is probabilistic, not certain. Because the conclusion is not based on a preconceived truth, but rather observations, we can be confident about the findings, but never certain.

Shifting to the evaluation context, one might wonder when deductive reasoning is best applied. When a commissioner has predetermined hypotheses that require answers is the obvious initial thought. While arguing against such an approach, Patton (2011) suggests interventions that have been designed through systematic applications of theory of change or logic models would lend themselves to deductive reasoning. Assuming the long-term outcomes of either have been met, starting with the hypothesis

that the intervention has had a causative effect provides a meaningful structure through which to collect data that supports or disproves the theory.

In instances where there is not a clearly articulated long-term outcome for a program, an inductive approach may be more appropriate. Starting with nothing more than a generalization, perhaps, *this intervention has had a positive effect*, an inductive approach would see the evaluator collecting comparably general and generic information about the intervention, with the intention of locating evidence which explains what the positive effect is, the magnitude of the effect and the mechanisms which caused it. However, there are situations that are less clear-cut. For example, a commissioner may order an evaluation that does not seek to validate an existing theory of change, in order to deduce if the expected changes and if the theory of change were sound, it should emerge from the inductive process anyway.

One of the societal shortcomings with reasoning is that inductive and deductive reasoning are often confused. There is a tendency to mistakenly label inductive reasoning as deductive reasoning. This is best exemplified through popular culture characters such as Sherlock Holmes, who is often labelled as having incredible powers of deduction. Holmes amazes audiences through his climactic reveals, which consist of numerous astute, seemingly unrelated observations that, when connected, provide a plausible (and incredible) explanation of the facts. This is only apparent to Holmes because he enters each situation with no preconceived notions. However, as demonstrated, these powers of observation that Holmes possesses are so impressive to audiences are most often demonstrations of inductive reasoning. He makes observations leading to generalizations or conclusions highly likely to be the explanation for the events that transpired (the approach is well illustrated in the BBC series *Sherlock*). If it were deductive reasoning, Holmes would have started with a clear hypothesis (whether the audience was aware at the time). Deductive approaches do exist in the Holmes universe, but are presented through the juxtaposing characters of John Watson and Detective Lestrade. If this was not compelling enough, the literary Holmes makes his position about his approach to reasoning clear on multiple occasions but most memorably in *A Study in Scarlet*, "it is a capital mistake to theorize before you have all the evidence. It biases the judgement" (Doyle 2007, xiv). As such, there is clearly a popular view that deductive reasoning is something to be proud of and amazed by. However, what this illustrates again is how these important philosophical choices are poorly understood, yet they make a dramatic impact on the way a problem is addressed or solved.

Disappointingly, some scholars tend to conflate deductive or confirmatory approaches with quantitative methods and deduction and induction exclusively (Creswell and Clark 2017, as cited in Soiferman (2010)). Trochim and Donnelly (2006) rightfully argue that this is an ill-informed generalization, and qualitative and quantitative methods should be in no way associated with deduction or induction. The vast majority of methods are appropriate for both approaches; it is what and how the methods are being used that determines whether the approach is appropriate.

5 | The Role of Abductive Reasoning

Finally, it is important to discuss the role of abductive reasoning in decision-making. While abductive reasoning does not operate on the deductive-inductive spectrum, it does play an important role in how decisions are made in daily life. Abductive reasoning was developed by American Charles Sanders Peirce in the late 1800s (Lipton 2017). Further developed by Harman (1965, 88), it can be summarized as "inference to the best explanation." It requires two prerequisites to be met in order for it to be relevant. First, there must be incomplete or the suspicion of incomplete data available to the decision maker. Second, there needs to be multiple, possible conclusions for the data that is available. This is where abductive reasoning departs from deductive and inductive reasoning. Rather than arriving at one conclusion, multiple conclusions are reached, and the preferred one is based on likelihood (Lipton 2017).

Sandoval-Hernández and Rutkowski (2024, 9) explains that abductive reasoning is most useful in making sense of "complex, ambiguous situations…where decision-makers must navigate uncertainty and incomplete information." Physicians are trained to use abductive reasoning through the differential diagnosis process (Magnani 1992; Martini 2023; Wackerly et al. 2024). Consider the last time you had a lingering cough and sought the advice of your physician. You might have been concerned you had a bacterial infection requiring antibiotics, or heaven forbid, something catastrophic like a pulmonary embolism or lung cancer. Physicians like you understand that there are multiple explanations for symptoms. However, their training provides them with the skill to attribute likelihoods to the various possibilities with great accuracy. This is why you are often reassured that the cough is merely a lingering artifact of a viral condition and nothing more sinister. This is also why, in some cases, it takes a significant amount of time to receive a devastating diagnosis, as the physician methodically works through the most likely possibilities. While this might seem counterintuitive, this often results in better outcomes for everyone. The patient is quickly reassured, is not subject to costly and invasive investigative procedures, and it frees up the health system to accommodate individuals with higher probabilities of serious outcomes.

For popular culture fans, the TV show *House MD* provides a good illustration of how this process takes place (Shore 2004–2012). For context, Dr House (who the show is named after) is a world-leading diagnostician. He and his team take on the cases that other physicians have been unable to solve, meaning all possibilities discussed are exceedingly rare, but discussions of likelihood relative still take place. A whiteboard is frequently used to communicate to the audience how the list of symptoms is then translated into possible diagnoses. After a spirited discussion, the team settles on the most likely explanation at the time, starting treatment and testing to confirm.

Translated to the evaluation space, abductive reasoning often has a strong role to play toward in intervention design as well as the concluding phases of an evaluation. As Harris et al. (2025) document, the intervention (re)design and conclusion of an evaluation often overlap when one considers the complete and ongoing evaluation cycle. Sandoval-Hernández and Rutkowski

(2024) provide an example during the intervention design phase. Here, a school principal generates several possible theories for decreasing mathematics results. The principal is acting on incomplete information but develops a theory they believe to be the most plausible at the time. From this point, one can take a deductive approach to test this theory. It may turn out that the deductive process demonstrates the theory is incorrect. However, it allows for action and progress even when faced with uncertainty. As the examples from medicine above indicate, once the first theory fails, one moves onto the next most likely explanation.

At the conclusion of an evaluation, the abduction process will look different. It is difficult for abductive reasoning to take place following a deductive approach. Because deduction ultimately results in the confirmation or nullification of a hypothesis, should it be nullified, one has to start the process again. There is no alternative to assess. However, as we have explained, because induction works from specifics in order to reach a generalization, evaluators may find during their data collection that multiple and competing generalizations emerge. This is not a failing of the inductive approach, it merely demonstrates that real-world interventions are complex. Outcomes and the mechanisms for those outcomes could plausibly be explained by multiple theories. Moreover, it is possible that the interaction of multiple theories accounts for the change. Either way, abductive reasoning provides a logical solution by pointing to the most probable explanation, thus allowing for further testing and future innovation.

Because abductive reasoning provides a logical solution for multiple possibilities in the face of incomplete data, evaluators need to be aware of this and intentionally apply it to their conclusions. Moreover, this abductive approach should also be specified in final reporting so readers understand how and why it is appropriate for some interventions to have multiple explanations of impact.

6 | Bringing Context, Philosophical, and Reasoning Decision-Making Into Methodological Choice

This article has outlined why it is important for evaluators to understand how context, philosophical and reasoning approaches are important information that informs the selection of methodologies leading to a rigorous evaluation design. In Figure 1 below, we present a decision-tree flowchart that outlines how understanding the context is crucial to identifying whether to adopt a holistic (where data gathering informs the outcomes) or reductionist approach (where there are highly specific questions to answer). In turn, the next decision is whether to adopt a deductive (a preconceived idea of what is happening is known) or an inductive approach (where no preconceived idea of what is happening is known). Ultimately, the inductive approach may lead to a series of outcomes where abductive reasoning (where one outcome is more likely than the others to have impact) comes into play. From this level of understanding and decision-making, the evaluator is now prepared to select one or more methodologies to build into an evaluation design.

For example, an evaluation design that required a reductionist/deductive approach was launched in June 2024 by the Paul Ramsey Foundation (Paul Ramsay Foundation 2024). With AUD$2.1 million in grant funding, the foundation was interested in awarding seven experimental evaluation designs, with each to receive up to AUD$300,000. At this point in time, it is unknown as to whether there were seven randomized control trials for social programs available for experimental evaluation. That aside, this is an example of a weak signal for change from inductive to deductive evaluation design.

A second example, another evaluation commissioned by the Australian Department of Social Services (2017) commissioned a holistic/inductive approach to evaluation of 52 pilot projects. This grant round encouraged new suppliers from corporate, nonprofit, and government agencies to partner together and develop innovative and new programs to address social disadvantage in young parents, young carers, students, mature unemployed, and youth at-risk. The Request for Quotation was deliberately general on the evaluation requirements and encouraged evaluators to be innovative and creative in their evaluation design across the entire project and at the individual pilot program level. This is an example of Australia's investment in pilots to determine whether an individual program was worthy of advancing through to whole-scale implementation across the country.

It is important to remind ourselves that the commissioners of evaluation set the parameters in which an evaluation will be conducted. This may be through a 30–45-min meeting where the decisions within the *Philosophical and Reasoning Evaluation Decision Tree* (Figure 1) are unpacked and addressed, or this could be through a published Request for Quotation (including subsequent questions and answers in the tender process). To strengthen the evaluation design, it is imperative that both commissioners and evaluators are clear about their positioning, leading to the selection of methodologies.

7 | Conclusion

In many respects, there is nothing revolutionary presented in this article. The philosophical debates regarding holism and reductionism, and deductive and inductive reasoning have existed for centuries. However, this article has shed light on these concepts in the evaluation context, which have historically avoided significant discussion. As detailed above, many of the evaluation's most revered thinkers have commented on these debates, but often very implicitly. We believe that evaluators are instinctively making these decisions, but perhaps the language used in this article is not in our evaluation lexicon. Owing to the multidisciplinary and highly contextual nature of evaluation, none of the positions are necessarily gold standard for evaluation, and we do not suggest any should be preferred over others. Rather, it is expected that this article will demonstrate to evaluators why it is important to be deliberate and aware of these choices early in the evaluation process, and the potential consequences it will have on the conclusions. Ultimately, there is a reasonable expectation that commissioners of evaluation, be they government entities and (where delegated) non-profit organizations, do need to be aware of the criticality of understanding the context, philosophical and reasoning approaches they require to assist evaluators to apply the most appropriate methodologies to enhance the evaluation outcomes.

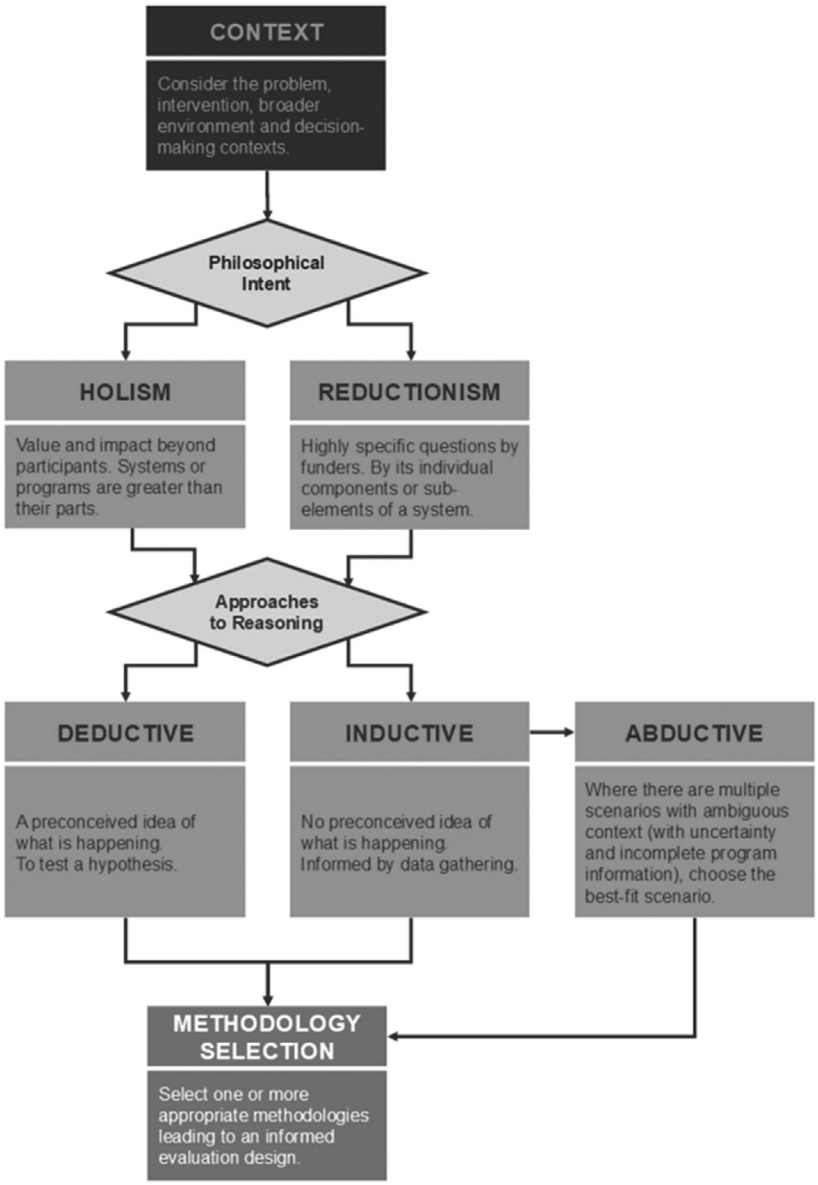

FIGURE 1 | Philosophical and reasoning evaluation decision tree.

Acknowledgments

Open access publishing facilitated by University of Southern Queensland, as part of the Wiley - University of Southern Queensland agreement via the Council of Australian University Librarians.

References

Ackoff, R. L. 1974. *Redesigning the Future: A Systems Approach to Societal Problems*. John Wiley & Sons.

Alderman, L., and B. Harris. 2025. "Editors' Notes." *New Directions for Evaluation* 2025: 185–186.

Australian Department of Social Services. 2017. Try, Test and Learn Fund. https://www.dss.gov.au/try-test-and-learn.

Bray, J. R., M. C. Gray, and D. I. Stanton. 2020. "Performance Management and Evaluation Meets Culture and Politics: Australia's Experience." *Asia Pacific Journal of Public Administration* 42, no. 4: 290–302. https://doi.org/10.1080/23276665.2020.1808491.

Buckley, J., T. Archibald, M. Hargraves, and W. M. Trochim. 2015. "Defining and Teaching Evaluative Thinking: Insights From Research on Critical Thinking." *American Journal of Evaluation* 36, no. 3: 375–388. https://doi.org/10.1177/1098214015581706.

Caruana, L. 2005. "God's Eternity and Einstein's Special Theory of Relativity." *Revista Portuguesa De Filosofia* 61, no. 1: 17–37. http://www.jstor.org/stable/40338163.

Coldwell, M., and N. Moore. 2024. "Learning From Failure: A Context-Informed Perspective on RCTs." *British Educational Research Journal* 50, no. 3: 1043–1063.

Cole, M. J. 2023. "Evaluative Thinking." *Evaluation Journal of Australasia* 23, no. 2: 70–90. https://doi.org/10.1177/1035719X231163932.

Conner, R. F., J. L. Fitzpatrick, and D. J. Rog. 2012. "A First Step Forward: Context Assessment." *New Directions for Evaluation* 2012, no. 135: 89–105. https://doi.org/10.1002/ev.20029.

Correll, M., E. Alexander, D. Albers, A. Sarikaya, and M. Gleicher. 2014. "Navigating Reductionism and Holism in Evaluation." In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Meth-*

ods for Visualization, 23–26. https://dl.acm.org/doi/abs/10.1145/2669557.2669577.

Creswell, J. W., and V. L. P. Clark. 2017. *Designing and Conducting Mixed Methods Research*. 3rd ed. SAGE Publications Inc.

Doyle, A. C. 2007. *A Study in Scarlet*. Random House Publishing.

Gensler, H. 2010. *Introduction to Logic*. 2nd ed. Routledge. https://doi.org/10.4324/9780203855003.

Greene, J. 2005. "Context." In *Encyclopedia of evaluation*, edited by S. Mathison, Sage.

Harman, G. 1965. "The Inference to the Best Explanation." *The Philosophical Review* 74, no. 1: 88–95. https://doi.org/10.2307/2183532.

Harris, B., L. Alderman, and J. Staheli. 2025. The Forgotten Contexts of Evaluation Evaluation. https://doi.org/10.1177/13563890241312910.

Hayes, B., and E. Heit. 2018. "Inductive Reasoning 2.0." *WIREs Cognitive Science* 9: e1459. https://doi.org/10.1002/wcs.1459.

Heit, E. 2007. "What Is Induction and Why Study It?" In *Inductive Reasoning*, edited by A. Feeney and E. Heit, 1–24. Cambridge University Press.

Horne, C. S. 2017. "Assessing and Strengthening Evidence-Based Program Registries' Usefulness for Social Service Program Replication and Adaptation." *Evaluation Review* 41, no. 5: 407–435. https://doi.org/10.1177/0193841X15625014.

House, E. R. 1993. *Professional Evaluation: Social Impact and Political Consequences*. Sage Publications.

Hume, D. 2000. *An Enquiry Concerning Human Understanding*. Edited by T. L. Beauchamp. Oxford University Press.

Hutchins, B. R. 2015. "Descartes, Corpuscles and Reductionism: Mechanism and Systems in Descartes' Physiology." *Philosophical Quarterly* 65, no. 261: 669–689. https://doi.org/10.1093/pq/pqv042.

Klauer, K. J., and G. D. Phye. 2008. "Inductive Reasoning: A Training Approach." *Review of Educational Research* 78, no. 1: 85–123. https://doi.org/10.3102/0034654307313402.

Laszlo, A., and S. Krippner. 1998. "Systems Theories: Their Origins, Foundations and Development." *Advances in Psychology* 126: 47–74. https://doi.org/10.1016/S0166-4115(98)80017-4.

Lemire, S., L. R. Peck, and A. Porowski. 2020. "The Growth of the Evaluation Tree in the Policy Analysis Forest: Recent Developments in Evaluation." *Policy Studies Journal* 48: S47–S70. https://doi.org/10.1111/psj.12387.

Lipton, P. 2017. "Inference to the Best explanation." In *A Companion to the Philosophy of Science*, edited by W. H. Newton-Smith, 184–193. Blackwell Publishers. https://doi.org/10.1002/9781405164481.ch29.

Magnani, L. 1992. "Abductive Reasoning: Philosophical and Educational Perspectives in Medicine." In *Advanced Models of Cognition for Medical Training and Practice*, edited by D. A. Evans and V. L. Patel, 21–41. Springer Science and Business Media.

Martini, C. 2023. "Abductive Reasoning in Clinical Diagnostics." In *Handbook of Abductive Cognition*, edited by L. Magnani, 467–479. Springer International Publishing. https://doi.org/10.1007/978-3-031-10135-9_13.

Mertens, D. 2024. "Visualizing Evaluation Theory: Tree, Forest, or Ocean Currents?" *Journal of MultiDisciplinary Evaluation* 20, no. 48: 17–22. https://doi.org/10.56645/jmde.v20i48.1015.

Midgley, G. 2006. "Systems Thinking for Evaluation." In *Systems Concepts in Evaluation: An Expert Anthology*, edited by B. Williams and I. Imam, 11–34. American Evaluation Association.

Newton-Levinson, A., M. Higdon, J. Sales, L. Gaydos, and R. Rochat. 2020. "Context Matters: Using Mixed Methods Timelines to Provide an Accessible and Integrated Visual for Complex Program Evaluation Data." *Evaluation and Program Planning* 80: 1–10. https://doi.org/10.1016/j.evalprogplan.2020.101784.

Patton, M. Q. 1978. *Utilization-Focused Evaluation*. Sage Publications.

Patton, M. Q. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. Guilford Press.

Patton, M. Q. 2018. "A Historical Perspective on the Evolution of Evaluative Thinking." *New Directions for Evaluation* 2018, no. 158: 11–28. https://doi.org/10.1002/ev.20324.

Patton, M. Q. 2019. "Expanding Futuring Foresight Through Evaluative Thinking." *World Futures Review* 11, no. 4: 296–307. https://doi.org/10.1177/1946756719882910.

Patton, M. Q., and C. E. Campbell-Patton. 2024. "Using Principles to Guide Evaluation Theory and Practice."In *Research Handbook on Program Evaluation*, edited by S. I. Donaldson and C. A. Christie, 137–160. Edward Elgar Publishing.

Paul Ramsay Foundation. 2024. Experimental Evaluation Design Open Grant Round. https://www.paulramsayfoundation.org.au/news-resources/experimental-evaluation-open-grant-round.

Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. Sage.

Raveh, E., Y. Ofek, R. Bekkerman, and H. Cohen. 2020. "Applying Big Data Visualization to Detect Trends in 30 Years of Performance Reports." *Evaluation* 26, no. 4: 516–540. https://doi.org/10.1177/1356389020905322.

Renn, M., and H. Gutfreund. 2023. *The Einsteinian Revolution: The Historical Roots of His Breakthroughs*. Princeton University Press.

Rog, D. J. 2012. "When Background Becomes Foreground: Toward Context-Sensitive Evaluation Practice." *New Directions for Evaluation* 2012, no. 135: 25–40. https://doi.org/10.1002/ev.20025.

Sandoval-Hernández, A., and D. J. Rutkowski. 2024. "Embracing Complexity: Abductive Reasoning as a Versatile Tool for Analyzing International Large-Scale Assessments." Education, Assessment, Evaluation and Accountability. https://doi.org/10.1007/s11092-024-09449-2.

Schechter, J. 2013. "Deductive Reasoning." In *The Encyclopedia of the Mind*, edited by H. Pashler. SAGE Reference.

Scriven, M. 1991. "Prose and Cons About Goal-Free Evaluation." *Evaluation Practice* 12, no. 1: 55–62. https://doi.org/10.1177/109821409101200108.

Scriven, M. 1994. "The Final Synthesis." *Evaluation Practice* 15, no. 3: 367–382. https://doi.org/10.1177/109821409401500317.

Senge, P. 1990. *The Fifth Discipline: The Art & Practice of the Learning Organization*. Currency.

Senge, P. 1997. "The Fifth Discipline." *Measuring Business Excellence* 1, no. 3: 46–51. https://doi.org/10.1108/eb025496.

Shanahan, T. 1989. "Deductive and Inductive Arguments." In *Internet Encyclopedia of Philosophy*, edited by J. Fieser and B. Dowden. https://iep.utm.edu/deductive-inductive-arguments/.

Shore, D., Executive Producer. 2004–2012. House, M.D. [TV Series]. Heel and Toe Films; Shore Z Productions; Bad Hat Harry Productions; Universal Television.

Smuts, J. C. 1926. *Holism and Evolution*. Macmillan.

Soiferman, K. 2010. Compare and Contrast Inductive and Deductive Research Approaches [Unpublished manuscript]. University of Manitoba.

Stufflebeam, D. L. 1968. *Evaluation as Enlightenment for Decision Making*. Ohio State University Evaluation Center.

Thomas, C. G. 2021. "Reasoning in Research." In *Research Methodology and Scientific Writing*. 2nd ed. Springer. https://doi.org/10.1007/978-3-030-64865-7.

Thomas, D. R. 2006. "A General Inductive Approach for Analyzing Qualitative Evaluation Data." *American Journal of Evaluation* 27, no. 2: 237–246. https://doi.org/10.1177/1098214005283748.

Trochim, W., and J. Donnelly. 2006. *The Research Methods Knowledge Base*. Cengage Learning.

Verschuren, P. J. M. 2001. "Holism Versus Reductionism in Modern Social Science Research." *Quality & Quantity* 35: 389–405. https://doi.org/10.1023/A:1012242620544.

Vo, A. T., J. S. Schreiber, and A. Martin. 2018. "Toward a Conceptual Understanding of Evaluative Thinking." In *Evaluative Thinking*, edited by A. T. Vo and T. Archibald, 29–47. New Directions for Evaluation. 158.

Wackerly, J. W., M. T. Wentzel, and S. K. Zingales. 2024. "Teaching Abductive Reasoning for Use as a Problem-Solving Tool in Organic Chemistry and Beyond." *Frontiers in Education* 9: 00. https://doi.org/10.3389/feduc.2024.1412417.

Weiss, C. H. 1998. *Evaluation: Methods for Studying Programs and Policies*. Prentice Hall.

Williamson, T. 2002. *Knowledge and Its Limits*. Oxford University Press.

Wilson, D. C. 2020. *A Guide to Good Reasoning: Cultivating Intellectual Virtues*. University of Minnesota Libraries Publishing.

Wittgenstein, L. 1958. *Philosophical Investigations*. 2nd ed. Macmillan Publishing.