A FRAMEWORK FOR

GO/NO-GO DECISIONS

IN HIGH-TECH R&D

A thesis

submitted partial fulfilment

of the requirements for the degree

of

Master of Management Studies

at

The University of Waikato

by

OWEN WOOLLASTON

THE UNIVERSITY OF

WAIKATO

Te Whare Wānanga o Waikato

Year of submission: 2025

Originality, Attributions and License

Declaration

This thesis was written in accordance with the University of Waikato Regulations and Guidelines.

This author has no conflicts of interest to disclose.

Use of Generative AI

Generative AI (GPT-4o) was used for the following tasks:

1. An AI-assisted literature search, as described in the method section.
2. Extracting and generating the terminology, abbreviations and acronyms.
3. Comment generation to describe software code behaviour in the appendices.
4. Ad hoc searches, often to identify key studies in a particular field.

Grammarly was used to review spelling and grammar.

Open Source Software Attributions

The following open-source software for the AI-assisted search:

1. Python 3.10.13 – general programming language
2. Openai 1.17.0 – Python-based library to access OpenAI (GPT).
3. Pybliometrics 4.0 – Python-based library to access Scopus
4. Unidecode 1.3.8 – ASCII transliterations of Unicode text

Image Attributions

All the figures are original works except for Figure 1 which contains graphics sourced from Canva.com under a paid license.

License

Correspondence concerning this thesis should be addressed to owen.woollaston@gmail.com.

Acknowledgements

I would like to thank my work colleagues, particularly my peers in the senior team, for their patience, understanding, and support. Bridget, you picked up the slack when I needed to focus on this work; Dean, you encouraged excellence; Mark, you helped spark this topic into life; and Sam, you stepped up at the last minute to review my writing.

To my supervisors, Dr Jenny Gibb and Prof. Frank Scrimgeour, thank you for your encouragement and pragmatic, experienced guidance in completing this. If it weren't for you, I might still be chasing rabbits down holes.

To my wife, Nicky, and children, Emily, Madeline, Liam, and Fin, thank you for covering for me in a thousand ways and supporting me despite losing family time. Without your love and support, this would not have happened.

Abstract

New product development (NPD) failure rates range from 40% (Cooper, 2019) to 90% (Kim et al., 2016), so firms must allocate resources effectively, including abandoning failing projects if necessary. However, when to abandon is often unclear, so portfolio managers risk commitment escalation (Eliens et al., 2018) and overinvesting. This study explored the research question: *How should firms evaluate whether to continue or abandon projects in high-tech R&D?* A dual approach was taken: First, to gain foundational insights, an innovative AI-assisted literature review method was developed and used to create a novel R&D decision-making framework. Second, to test the key concepts of the framework and illustrate its application, an agent-based simulation (Sulis & Taveter, 2022) compared the performance of the framework against risk-based and ROI-based approaches in a 1000-project portfolio. The framework outperformed the other strategies in simulations and emphasised the importance of gatekeeper independence, negotiating clear criteria up front and using formal decision gates. The simulation results also suggested actionable guidance for practitioners: For a feasibility phase to add value, it should cost less than 30% of the project budget and significantly improve cost estimates.

Theoretical contribution: This study integrated key NPD decision theories from real options, behavioural decision-making, and portfolio management and developed a novel decision-making framework while exemplifying the use of AI to enable rapid, scalable research.

It then demonstrated that decision gates outperform ROI- and risk-based approaches in simulation.

Practical contribution: For R&D portfolio managers, the study provides a structured decision-making approach and guidance for sizing feasibility studies. For researchers, it provides a methodology for accelerating and scaling literature reviews.

Keywords

AI-assisted search, new product development, R&D portfolio management, R&D decision-making, R&D project management, stage-gate.

Table of Contents

Table of Figures

Table of Tables

Terminology, Abbreviations and Acronyms

Agent-Based Modeling (ABM) – A computational simulation approach that models the interactions of autonomous agents to assess their collective effects on a system.

Antecedents – Pre-existing factors, such as organisational strategy or capabilities, that influence decision-making processes.

Availability Heuristic – A cognitive bias in which individuals estimate probabilities based on how easily examples come to mind.

Bias – Systematic deviations from rational judgment due to cognitive limitations, emotional influences, or decision-making shortcuts.

Bounded Rationality – A concept suggesting that decision-makers operate under cognitive and informational constraints, leading them to settle for satisfactory solutions rather than optimal ones.

Business Case Gate – A formal decision checkpoint where the feasibility and expected returns of a project are evaluated before further investment.

Cognitive Biases – Systematic errors in human thinking that can lead to suboptimal decision-making.

Commitment Escalation – The tendency to continue investing in a failing project due to past investments and psychological biases.

Concept Gate – The initial screening point in an NPD process where new project ideas are assessed for feasibility.

Critical Realism – A philosophical perspective that acknowledges that perceived reality may differ from objective reality, influencing decision-making.

Decision Gate – A structured evaluation point in an NPD process where a go/no-go decision is made based on available information.

Decision Tree Analysis – A graphical representation of possible choices, risks, and outcomes to support structured decision-making.

Discrete Event Simulation (DES) – A modeling technique where system state changes occur at specific event-driven points rather than continuously.

Distancing Strategy – A decision-making approach that reduces bias by involving independent evaluators or external perspectives.

Dynamic Capabilities – The ability of an organisation to adapt and reconfigure its resources in response to changing market conditions.

Feasibility Study – A detailed analysis that assesses the practicality and financial viability of a proposed project.

Forecasting Error – The deviation between predicted and actual project outcomes due to uncertainty in market or cost estimations.

Game Theory – A mathematical framework for modeling strategic interactions between competing decision-makers.

Gatekeeper Independence – The principle that decision-makers evaluating a project should not be personally invested in its success to maintain objectivity.

Go/No-Go Decision – A binary decision-making process determining whether a project should proceed to the next phase or be terminated.

Heuristics – Mental shortcuts or rules of thumb that simplify decision-making but can introduce biases.

Innovation Funnel – A structured process that filters potential R&D projects through multiple decision points to optimise resource allocation.

Internal Rate of Return (IRR) – A financial metric used to assess the expected profitability of an investment.

Judgment Under Uncertainty – The study of how individuals make decisions when probabilities and risks are unknown.

Knightian Uncertainty – A form of uncertainty where the probability of outcomes is entirely unknown and cannot be quantified.

Large Language Models (LLMs) – AI models trained on vast datasets to generate human-like text and assist in tasks like research and decision support.

Loss Aversion – A cognitive bias in which individuals weigh potential losses more heavily than equivalent gains.

Market Volatility – Fluctuations in market conditions that introduce uncertainty into financial and investment decisions.

Monte Carlo Simulation – A statistical method that uses repeated random sampling to model uncertainty in financial and decision-making processes.

Net Present Value (NPV) – A financial metric that calculates the present value of an investment by discounting future cash flows.

New Product Development (NPD): The structured process of applying R&D outputs and other inputs to design, develop, and commercialise new products, typically with clearer market objectives and governance controls.

Opportunity Cost – The cost of choosing one investment alternative over another, considering the potential benefits foregone.

Option Analysis – A decision-making approach that evaluates different strategic choices, often used in real options theory.

Portfolio Management – The process of selecting, prioritizing, and managing a set of R&D projects to align with strategic objectives.

Positivism – A philosophical approach in research that emphasizes objective measurement and observable data.

Project Governance – The framework of policies and decision-making structures that guide project execution and investment.

Real Options Theory – A financial and strategic decision-making framework that treats investment decisions as options that can be exercised under uncertainty.

R&D (Research and Development): Organisational activities focused on generating new knowledge, technologies, or capabilities, often under uncertainty, to enable future innovation opportunities.

Risk-Based Decision-Making – A decision-making approach that explicitly considers probability and potential negative consequences.

Scenario Analysis – A decision-support technique that evaluates potential outcomes based on different assumptions.

Sensitivity Analysis – A method used to determine how different input variables impact a given outcome.

Stage-Gate Process: A structured framework for managing new product development (NPD) projects, where progress is evaluated at predefined decision points (gates) to control risk and allocate resources effectively.

1 Introduction

1.1 Research Context and Motivation

New product development (NPD) has long been recognised as a driver of organisational and economic growth (Schumpeter, 1934; Nelson, 1961; Kohn et al., 2021). In modern firms, research and development (R&D)[1] are at the heart of innovation, and high-tech firms invest heavily in developing new products to maintain a competitive edge (Jin et al., 2024; Zaman & Tanewski, 2024). Most firms must choose how to allocate limited resources to R&D projects for maximum economic advantage in an environment where development costs and timeframe are uncertain, as are returns (Acebes et al., 2022). Effective decision-making and R&D portfolio management are central to making suitable R&D investments (Cooper et al., 1997; Cooper, 2022), and no decision is more fundamental to portfolio success than the decision to continue an R&D project or abandon it and repurpose resources elsewhere.

Despite the body of knowledge around R&D best practices, failure rates of R&D projects are generally acknowledged to be high – R&D is inherently risky. However, extensive studies that quantify failure rates are challenging to find. Crawford (1987) predicts failure rates to be around 35%. Boulding et al. (1997) quote R&D failure rates between 35 and 45%. Two decades later, Cooper (2019) reaffirmed, "About 40% of new products are estimated to fail at launch." Conversely, Kim et al. (2016) cite the failure rate in one Korean company as 87%. Interestingly, in none of these publications was the failure rate the subject of the study.

A widely cited but unverified claim is that NPD failure rates are around 85%, but in "Myths about new product failure rates", Castellion and Markham (2013) debunk this statistic as an urban myth and collate empirical studies from 1977 to 2009 that indicate 40% is more typical. The paucity of research is evident in their discovery of only 12 studies across all industries in this period. Castellion and Markham (2013) also point out that organisations with lower process maturity (typically small to medium enterprises, "SMEs") have higher failure rates (55%), whereas "best in class" organisations achieve a 38% failure rate. All this suggests that R&D is risky, but the failure rate depends on the industry and the firm's maturity. What can be done to help high-tech SMEs avoid and learn from hard-won lessons and mature more quickly? This question is the key motivation behind this thesis.

In the context of NPD, failure can take on many forms – financial, commercial/market, regulatory, strategic or technical. The point in time at which failure occurs can also vary greatly. Regulatory

[1] While often used interchangeably, R&D and new product development (NPD) refer to distinct stages of innovation. R&D encompasses exploratory activities aimed at generating new knowledge or capabilities, typically under high uncertainty. NPD, by contrast, focuses on the structured development and commercialisation of products, often using outputs from R&D as inputs into a more governed process.

failure occurs after development but before product release. In contrast, technical failure can occur any time after release (sometimes resulting in expensive product recalls), and market and financial failure cannot be conclusively determined until the product has been in the market for a time. Ford, Aubert, and Ryckewaert (2016) orient failure on the NPD process itself, considering abandoned projects as failures – failure of the *project* to deliver the anticipated *product*. Conversely, this thesis focuses on *product* failure – the failure of the product to deliver commercial outcomes – which occurs after release. Thus, for this thesis, failure is defined as commercial (market or financial) failure of the product in the market, that is, after development, regulatory certifications and release.

Given the high failure rates, firms must take care to allocate resources effectively. This need has given rise to two management disciplines: portfolio management and project management. *Portfolio management* is the investment of resources into a set of R&D projects to meet strategic objectives, typically involving profit and market share. Cooper et al. (1997) define it as "...a dynamic decision process, whereby a business's list of active new product (and R&D) projects is constantly updated and revised. In this process, new projects are evaluated, selected and prioritised; existing projects may be accelerated, abandoned or de-prioritized; and resources are allocated and reallocated to the active projects." It is the fundamental process by which R&D investments are made. These are inherently more challenging to evaluate than traditional investments due to high cost/benefit uncertainty, complexity, many degrees of freedom in the design and execution, and the volatility of target markets (Dhillon & Nguyen, 2021).

Once an investment decision has been made, it becomes the realm of *project management* to plan, monitor and control the project, typically utilising processes such as Agile, Scrum, or Waterfall. R&D project management processes typically reduce uncertainty using stages, phases, iterations, releases, gates or milestones. At each step in the process – whatever the specifics – the business confidence is expected to increase toward launch. Jagle (1999) states: "Phase models of the NPD process also allow the transformation of unmeasurable uncertainty into measurable risk." and Loch (2000) also found that: "An individual project cannot be managed toward milestones and product feature deliverables, but must be managed toward uncertainty reduction." While *R&D portfolio management* is about resource allocation, *R&D project management* is about managing risk and uncertainty.

1.2 Conceptual Overview of R&D Decision-Making

With failure defined as a commercial failure after release, it follows that the portfolio investment decisions made *during* an R&D project have a critical bearing on the product's success. To manage this, many firms have adopted or evolved processes to formalise investment decision-making in product development. These are usually called stage-gate processes (Cooper, 2022). These processes define several decision gates ("gates") and associated criteria and are administered by a governance group (a "product council") to control R&D investment spending at each step. The product council will be presented with several options, the most fundamental of which is the decision to continue or abandon a project – colloquially "go/no-go" (Cooper, 2006, 2022; Huchzermeier & Loch, 2001). However, there are typically intermediate options that involve adapting the project to new information. The specifics of each option will vary by project and typically have many degrees of freedom. While options exist theoretically at all points on the development timeline, the stage-gate approach simplifies decision-making by identifying junctures at which critical information is available. For example, after product ideas have been initially screened and ranked, managers can effectively choose the most viable new products for the research phase. Alternatively, when market research is completed and technical risks evaluated, managers can reassess the viability of continuing to the development phase. Thus, stage-gates processes, and more precisely, the gates, facilitate effective R&D decision-making by reducing options to a manageable set of choices at predefined points in the process.

Gates form a critical link between project and portfolio management, that is, between risk management and resource allocation (Cooper, 2022; Cooper & Sommer, 2023). To understand this, it is necessary to clarify what is meant by "stage-gate" and "gate." A stage-gate process can be any staged process punctuated by defined decision gates. Stage-gates are typically sequential but may also be iterative. They may apply to design, manufacturing, research, or product development (Cooper, 1990). The definition used here is: a staged project management process delineated by decision gates and governed by a product council to control R&D portfolio investments and resource allocation. By this definition, gates form a critical communication link and feedback loop between two business systems: (1) the planning and execution of the project (encapsulating risk, cost, and benefit), and (2) the management of the product portfolio (encapsulating governance, resource allocation, and project prioritisation) by the governance group (Ayala-Cruz, 2016). Thus, the most critical feature of a stage-gate process is the escalation of go/no-go decisions from the project team to the governance group, and this thesis focuses on these critical gates.

1.3 Key Challenges in R&D Decision Making

For R&D portfolio managers, these investment decisions are fraught with potential errors. As if abundant options and high uncertainty were not enough, human factors frequently distort the perception of cost/benefit, resulting in missed opportunities and overspending (Buehler et al., 2002). For this reason, R&D decision-making (particularly project selection) has been studied through many lenses. From the perspective of traditional economics, real options theory provides quantitative approaches to evaluating options – namely, continue, abandon, pivot, resize, and defer

(Huchzermeier & Loch, 2001), although variations in terminology arise. Behavioural psychology has uncovered dozens of heuristics – rules of thumb (Sherman & Corty, 1984; Gilbert-Saad et al., 2018) – and shows us the many dark paths to commitment escalation, "throwing good money after bad" (Staw, 1976), such as impression management and overconfidence bias. Behavioural economics blends economics and psychology to help us quantify human bias (Kahneman et al., 2021; Kahneman & Tversky, 1977). The management literature spans from strategy to execution, offering venture governance (Sahlman, 1990; Garg, 2020), portfolio management (Martinsuo, 2013), and project management (Schwalbe, 2009; Levitt, 2011). Even more mind-bending is that modern neuroscience tells us that consciousness is an illusion and that what we experience is a form of controlled hallucination (Seth & Bayne, 2022) – a "Bayesian best guess" at reality. Although each academic perspective provides valuable insight separately, together, researchers have created an overwhelming plethora of considerations with little structure or clarity for practitioners. How should a biased, confused and overwhelmed manager with a partial grip on reality living in a volatile, uncertain, complex and ambiguous world make sound investment decisions?

R&D decision-making has evolved to assimilate the thinking of the time. Early thinking focused on objective, analytical approaches and the application of economic theory to R&D, particularly financial options. Real options theory became popular because of its relative flexibility and was well-studied in the NPD context by the turn of the millennium (Huchzermeier & Loch, 2001; Loch & Kavadias, 2002). However, the rational economic perspective of real options failed to explain observed outcomes. Recently, the findings of behavioural economists have begun to be incorporated as researchers explore the impact of human bias in R&D decision-making (Nielsen et al., 2024; Ross et al., 2017; Vaculik et al., 2019). Killen et al. (2008) highlight the importance of selecting projects that align strategically. Each researcher focuses on a specific slice of the problem, such as the effect of CEO overconfidence (Lee et al., 2022), but this approach provides only narrow guidance. Research naturally emerges as a conversation over years or decades like a wavefront, with each new ripple an article inspired by those that came before it. However, like a long email chain, many people simply do not have time to start from the beginning. The problem is that there is too much information and a lack of cohesion between sources. An integrated decision-making framework is needed.

To address these challenges, this thesis is structured in two parts. The first comprehensively explores and synthesises the literature on R&D decision-making using an AI-assisted search. The second part uses computer simulation to illustrate the efficacy of different decision-making strategies in different commercial environments. This dual approach reinforces theoretical findings with quantitative data, ensuring a comprehensive understanding of theory and application.

1.4 The Opportunity With Large Language Model AI

Given the high volume of information, contemporary researchers face a similar problem. As the number of publications from many sources has increased over the past decades, researchers (limited by bounded rationality) have been forced to develop ways to narrow the scope of their literature searches early to avoid long delays in processing large quantities of publications (Cronin & George, 2023; Xiao & Watson, 2019). Saunders et al. (2022) advocate identifying relevant journals, selecting reputable sources, selecting highly cited journals, and narrowing the search using boolean operators and keywords to "right size" the results for human consumption. Unfortunately, these techniques serve as coarse proxies for relevance. Just because the topic of the journal is relevant doesn't mean all relevant articles have been published in that journal or others like it (Wagner et al., 2022; Joos et al., 2024). Even with contemporary research methods, the rapidly increasing volume of information presents an escalating likelihood that relevant information will be omitted (Mourão et al., 2020; Bolaños et al., 2024).

Unlike human researchers, AI has no such capacity limitation. The general availability of ChatGPT in November 2022 (OpenAI, 2024) made programmatic interaction with large language model ("LLM") AI accessible to mainstream researchers. LLMs intrinsically interpret information based on concept rather than syntax (Shani et al., 2023; Havlík, 2023). This presented an opportunity to explore incorporating AI into research methods as a secondary goal of this thesis. Tóth and Oldal (2022) and Wagner et al. (2022) were early researchers in applying AI to research methods, and this author also saw an opportunity to make the research process more effective and efficient. Specifically, the opportunity was to drastically widen the initial search to thousands of potentially relevant articles and then use AI to rate their relevance to the research question. LLMs are perfect for this task because the relevance rating is based on conceptual relevance rather than word matching. For these reasons, the decision was to pioneer an AI-assisted literature search method to facilitate the development of an integrated framework.

1.5 Rationale for Simulation

Establishing the theoretical foundations and developing a decision-making framework is insufficient to demonstrate the value of this work. Some level of demonstration and validation is required. Unfortunately, real-world testing of a decision-making framework has significant challenges. Projects are usually too expensive to run twice with control and experimental teams, and samples are too small for statistical validity (Davis et al., 2007). This process would take years and possibly involve longitudinal studies to demonstrate worth. Setting up A-B studies to compare the proposed framework with other decision-making approaches is possible in controlled simulations with human participants. However, these generally suffer from variability between participants, and it is difficult to get large enough samples for statistical validity. Case studies would provide qualitative feedback but, again, lack empirical validity. Given the scope of this thesis and the time available, these approaches are unrealistic and unlikely to validate the theoretical findings significantly.

Unburdened by real-world complexities, computer simulation presents a robust alternative. In their work "A Flexible Multi-Metric Bayesian Framework for Decision-Making in Phase II Multi-Arm Multi-Stage Studies," Dufault et al. (2023) successfully used simulation to trial a decision-making framework used in clinical trials. This study highlights the advantages of simulation in testing decision-making frameworks compared to real-world trials. Repeatability is usually perfect (Davis et al., 2007), allowing multiple decision-making approaches to be tested against the same portfolio of projects. The ability to learn by exploring – to run and re-run the simulation – is highly valuable when refining the simulation; the cost to re-run is negligible, and the results are practically instantaneous. Further, simulations allow researchers to isolate and control variables that might not be controllable in the real world (for example, illness or attrition). Finally, scalability is very high – in simulations, it is possible to test using a portfolio of hundreds of projects (Lyons et al., 2020). For these reasons, the decision to validate using simulation was made.

In summary, the introduction began with high failure rates in NPD and outlined the context of portfolio management. Then, a conceptual framework was introduced in which go/no-go decision-making is central to R&D project management. Next, it defined the problem: A lack of cohesion between many excellent sources presents an overwhelming amount of information for practitioners – R&D portfolio managers who must decide which project to invest in and which to abandon. Finally, it argued the case for using AI in literature searches and for simulation to test the theoretical findings.

1.6 Research Question and Objectives

Following a review of the literature on R&D project evaluation and decision-making, particularly in relation to project termination and abandonment, this thesis focuses on projects already in progress. Within this context, the following research question is posed:

RQ: How should firms evaluate whether to continue or abandon projects in high-tech R&D?

Knowing when to abandon a failing project is a critical skill for R&D portfolio managers. Abandoning an unfruitful line of research retains critical resources for other, more attractive projects. While failure at launch is certainly undesirable, abandoning a project during its early stages might be a smart move.

The primary objective of this research is to develop an integrated framework to facilitate go/no-go decisions in high-tech R&D in terms of relevant theoretical constructs (for example, antecedents and criteria). Specifically, this work aims to:

1. **Conduct an AI-assisted literature review** to identify existing theories related to go/no-go decision-making in high-tech R&D.
2. **Develop an integrated decision-making framework** based on insights from the literature, providing structured guidance for go/no-go decisions.
3. **Validate the framework through computer simulation**, illustrating how different decision strategies perform in various commercial environments.

Findings from this study will contribute to the academic literature that begins with an integrated literature review that is followed by developing a conceptual model focused on go/no-go decisions, as well as advancing the practice of research by pioneering a new search method. For practitioners, the framework will serve as a checklist for evaluating the quality of their decisions and, more generally, a framework for maturing their portfolio management process. In doing so, this author hopes that in some modest way, R&D resources will be better spent, industry and academic productivity will rise, and society as a whole will be better off.

With an AI-assisted literature review backed up by practical illustrations using simulation, this thesis addresses the gap between fragmented theoretical perspectives and the practical need for integrated, usable decision frameworks, offering a novel framework for improving go/no-go decision-making in high-tech R&D.

1.7 Structure of this Thesis

This thesis is divided into two parts. Part A, "Integrating 30 Years of Research on R&D Portfolio Decision-Making – An AI-Assisted Literature Review", is an integrative literature review (Cronin & George, 2023) that develops from the literature and uses an AI-assisted search approach capable of ranking thousands of abstracts for relevance to the research question. Part A culminates in presenting a novel go/no-go decision-making framework and discussing the results, implications, and conclusions.

Part B, "Simulating R&D Investment Decisions – An Agent-Based Approach", is a set of simulations in which three decision-making agents ("ROI-based", "risk-based", and "gated") each evaluate two different portfolios of projects (one modelling high uncertainty, based on a uniform distribution, and one modelling "typical" development and market environments based on a family of beta distributions). These simulations illustrate the efficacy of different decision-making approaches when evaluating a portfolio of R&D projects in different development and market environments.

Simulations are developed, results are presented and discussed, and conclusions are drawn, along with practical guidance for sizing feasibility investments.

The final integrated conclusion summarises insights from the literature review and simulation, offering recommendations for practitioners and avenues for future research.

As an epilogue, Appendix G offers empirical validation of the AI-assisted approach, and Appendix H offers a brief review of the novelty of the same.

2 Part A: Integrating 30 Years of Research on R&D Portfolio Decision-Making – An AI-Assisted Literature Review

This literature review examines three decades of research on R&D portfolio decision-making, integrating insights from multiple disciplines through an AI-assisted systematic review. The review begins by establishing the theoretical foundations underpinning decision-making in R&D. It then details the methodology used to conduct the integrative review, including the AI-assisted search strategy, eligibility criteria, and data extraction process, alongside a critical assessment of potential biases. The synthesis method is then outlined, leading to a structured integration of the findings. The results present the current state of the literature, key factors influencing go/no-go decision-making, and a novel decision-making framework. Finally, the review highlights gaps in the literature and concludes by summarising key insights before transitioning to Part B, which explores simulation-based decision modelling.

2.1 Theoretical Foundation

This section introduces the key theories underpinning R&D decision-making, outlining common decision-making models, theoretical perspectives of uncertainty and risk, and psychological and cognitive influences, particularly commitment escalation. It then presents the research gap, along with the case for a literature review and a new go/no-go decision-making framework. Finally, it concludes with theoretical constructs that will serve as a basis for the literature review.

2.1.1 Uncertainty and Risk in R&D

A major contributor to the high failure rates in R&D is the uncertainty and risk associated with developing novel products for unpredictable and volatile markets (Kohn et al., 2021; Wang et al., 2010). When describing uncertainty, Knight (1921) identified that probabilities are often not known. As Keynes (1937, p. 214) put it, ‘‘About [uncertainty] there is no scientific basis on which to form any calculable probability whatsoever. We simply do not know.’’ This is called Knightian Uncertainty (the absence of a known probability distribution). However, Knightian Uncertainty still confers information which can be used to infer risk – specifically, relevant variables about a system that may be known and their relationships to each other. Schrader et al. (1993) take this one step further by discerning unforeseeable uncertainty: ‘‘the inability to recognise and articulate variables and their functional relationships.’’ Both types are common in R&D, but it is sufficient here to define uncertainty in the general sense covering both types. Examples include not knowing what issues a product development team will strike or whether a chosen technology will fulfil customer requirements. Conversely, in this thesis, risk implies knowledge of the probability of a possible outcome – probabilistic risk. Cooper (1999) identifies many risk factors in NPD, such as development overruns, which can be significant but are usually bounded and somewhat predictable. Risk, by definition, is quantifiable and can be evaluated analytically (Farshchian & Heravi, 2018; Kettunen & Salo, 2017).

The distinction between risk and uncertainty is crucial because they are treated differently. At acceptable levels, risk can be managed through mitigation strategies within a project, whereas uncertainty requires strategic flexibility at the portfolio or firm level (Huchzermeier & Loch, 2001; Loch et al., 2006; Mavrotas & Makryvelios, 2021). This difference underscores the need for tools and frameworks that accommodate both probabilistic risk and more ambiguous forms of uncertainty when making R&D investment decisions (Kauffman et al., 2015; Shakhsi-Niaei et al., 2011).

While new product developments in some industries follow predictable patterns (simple toys or food products, for example), this is rarely the case in high-tech. At the outset, firms often face uncertainty in technical feasibility, a high level of risk in the development efforts, and high risk or even uncertainty in the market returns. The process of R&D – and even the label "research and development" – embodies the notion of transforming uncertainty into manageable risk (Loch et al., 2006; Sommer et al., 2009). Here, manageable risk means a level of risk palatable to the firm's risk appetite. NPD's fundamental purpose is to take something as ephemeral as an idea and transform it into something highly reproducible and saleable (be that software services or physical products) (Cooper, 2022). The transformational process requires a scrupulous understanding of the project risks and uncertainty at every step for effective decision-making, and many frameworks have been developed to support this transformation.

2.1.2 Decision-Making in High-Tech R&D

Decision-making has been studied for decades. In the context of decision-making in firms, early economists developed game theory (Von Neumann & Morgenstern, 1944) and expected utility theory (Savage, 1954). However, these classical economic models failed to accurately describe real-world behaviour, so researchers quickly began acknowledging cognitive limitations, beginning with bounded rationality (Simon, 1955). Simon (1956) then formalised "satisficing," in which decision-makers settle for a good enough option rather than seeking the optimal one. In A Behavioral Theory of the Firm (Cyert & March, 1963), the authors demonstrate the use of satisficing, showing that firms often use rules of thumb (heuristics) rather than seeking optimal solutions. From there, cognitive and psychological influences in decision-making became more relevant in the seventies with Daniel Kahneman and Amos Tversky publishing in three consecutive years on judgement and heuristics – "Subjective probability: A judgment of representativeness." (Kahneman & Tversky, 1972), "Availability: A heuristic for judging frequency and probability." (Tversky & Kahneman, 1973) and

“Judgment under uncertainty: Heuristics and biases.” (Tversky & Kahneman, 1974) – and then their seminal work on prospect theory in 1979 (Kahneman & Tversky, 1979) which demonstrated asymmetry in human perception of losses vs gains. Prospect theory is particularly relevant when considering go/no-go decisions because it tells us that managers see abandoning a project as more significant than resuming another of equivalent value. These behavioural economists all highlight and explore the limitations in capacity (bounded rationality, satisficing) and accuracy of perception (bias and noise) that affect contemporary decision-makers and form the foundation of our understanding of human decision-making today (Kahneman et al., 2021; Del Campo et al., 2016; Gilbert-Saad et al., 2018; Roeth et al., 2019; Yang et al., 2020).

Satisficing, heuristics and biases are just as relevant in R&D as in any other field. Perhaps more so considering that R&D is typified by high uncertainty, particularly regarding the expected development effort and market returns. Decision-making, particularly the effective allocation of resources, is critical to success in NPD. Yet, uncertainty, ambiguity (in product requirements), the volatility of target markets, and the high cost of market research make decision-making complex and force decision-makers to rely heavily on intuition, heuristics and scant data (Del Campo et al., 2016).

2.1.2.1 Portfolio Management Frameworks

In response to the many challenges in R&D decision-making, academics and practitioners have developed various portfolio management tools, processes, and frameworks to facilitate new product development. There are as many portfolio management approaches as there are philosophical dispositions. At the rational end of the spectrum is value-based portfolio management (Koller et al., 2010; Cooper & Sommer, 2023) that leans heavily on numerical evaluations of economic value (for example, net present value) and the Stage-Gate® process (Cooper, 1990; Cooper, 2022) that assume decision-making follows a logical flow where objective data leads rational choice. The Boston Consulting Group (BCG) Matrix (Henderson, 1970; Chiu & Lin, 2019) takes an empirical perspective, assuming that markets are predictable and quantifiable based on past data. Real options theory (Trigeorgis, 1996; Posen et al., 2018) takes a more pragmatic perspective, encouraging retention of strategic options and openness to new possibilities, as do strategic buckets (Cooper et al., 1997; Hutchison-Krupat & Kavadias, 2015) that encourage firms to divvy up resources to ensure that breakthrough, sustaining and incremental innovations are adequately serviced. Finally, balanced scorecard (Kaplan & Norton, 1992; Hasan & Chyi, 2017) embraces critical realism by acknowledging that multiple perspectives (financial, customer, process and learning) are valuable.

Regardless of philosophical disposition, each portfolio management approach involves investment decision-making at critical points during an R&D project. Some, like the BCG Matrix, value-based portfolio management, strategic buckets and balanced scorecard, apply more generally to the portfolio as a whole and don't define project stages or gates. Thus, they inform the overall portfolio strategy – to define which technologies and markets to focus on, how to allocate (coarse) budgets, and how to predict and measure financial outcomes. They are only loosely coupled to the project decision-making under the "strategic alignment" category and serve as guidance to the delivery team. Further, they provide no specific guidance regarding go/no-go decision-making.

More relevant to this thesis, real options (Trigeorgis, 1996; Posen et al., 2018) and stage-gates (Cooper, 1990; Cooper, 2022) include the concept of mid-project stages and gates – formalised points at which a go/no-go decision is made. Real options make explicit the concept of identifying options to be evaluated at predefined moments during an R&D project. Thus, they are tightly coupled to the go/no-go decisions during the execution of a project. They are intrusive in that they serve as project management checkpoints for the health of the project and the future value of the investment. This interface, connecting project and portfolio management through go/no-go decisions, is where this thesis focuses and seeks to define and elaborate how to make these decisions effectively in service of broader portfolio and strategic objectives.

This narrow focus makes this thesis, at its core, process-agnostic. The focus is on the decision point at which the go/no-go decision is made, not the surrounding process. Thus, regardless of the process, the terms gate, decision gate and decision point are used synonymously to mean the point where a go/no-go decision is made. This does not mean that the specifics of each process are irrelevant or do not add value. As stated in the introduction, this thesis adopts a generalised definition of stage-gate, a staged project management process delineated by decision gates and governed by a management group to control R&D portfolio investments and resource allocation.

Acknowledging the contribution of the portfolio management approaches above, one theory – real options – closely aligns with this thesis with its explicit focus on decision-making during R&D projects and its philosophy of retaining options under uncertainty. In contrast, other frameworks offer only strategic guidance for mid-project decision-making.

2.1.2.2 Real Options Theory and Gates in R&D

Having strategic options is critical in R&D. Often, things do not go according to plan, so R&D processes have evolved to identify key decision points, but this can make them complex to describe. In Cooper's Stage-Gate® process, there are six stages: Discovery, scoping, business case development, development, testing and validation, and launch, and there are five gates (no/no-go decisions) separating each of the six stages: Idea screen, second screen, go to development, go to testing, go to launch (Cooper, 2022). This is more detail than is conceptually necessary in this thesis, so to simplify, a generic three-stage process is assumed:

1. **Selection** – Generation, screening, and prioritisation of early concepts.
2. **Feasibility** – Exploration of the concept's commercial and technical feasibility to translate uncertainty into a manageable (palatable) risk level.
3. **Development** – The product development, which may be broken into milestones. The finished product is the result of this phase, and the gate is the decision about whether and how to launch the product.

The point is that there is a process, and at each step, there are natural exit points, which are go/no-go gates. Whether the process has three stages or six does not matter here. What matters is the process is a sequence of stages punctuated by gates. This representation of the process makes R&D investment a series of real options rather than a fixed up-front commitment. This structure is highly beneficial because it allows managers to minimise investment in failing projects through early detection at pre-programmed points in the project.

Figure 1

Gates and Options in R&D

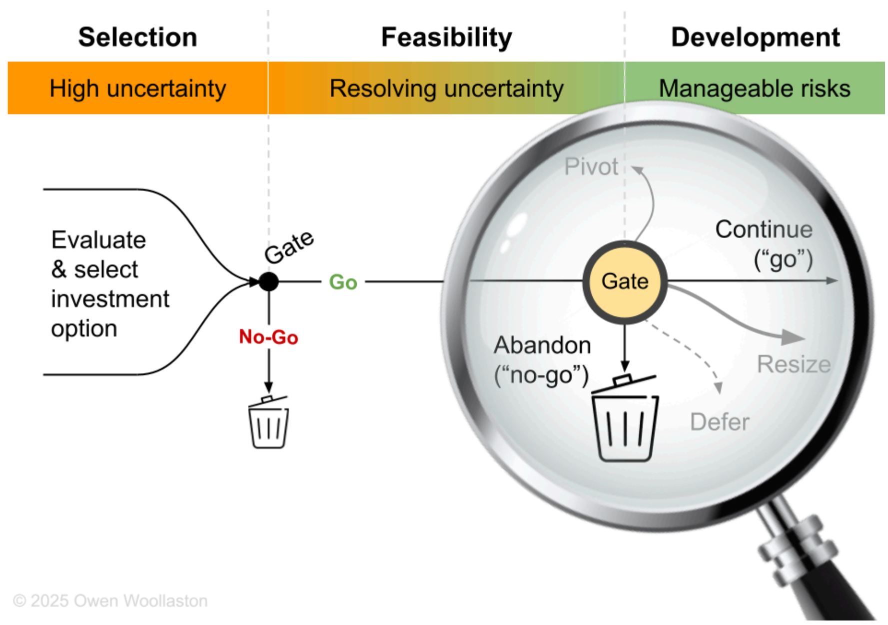
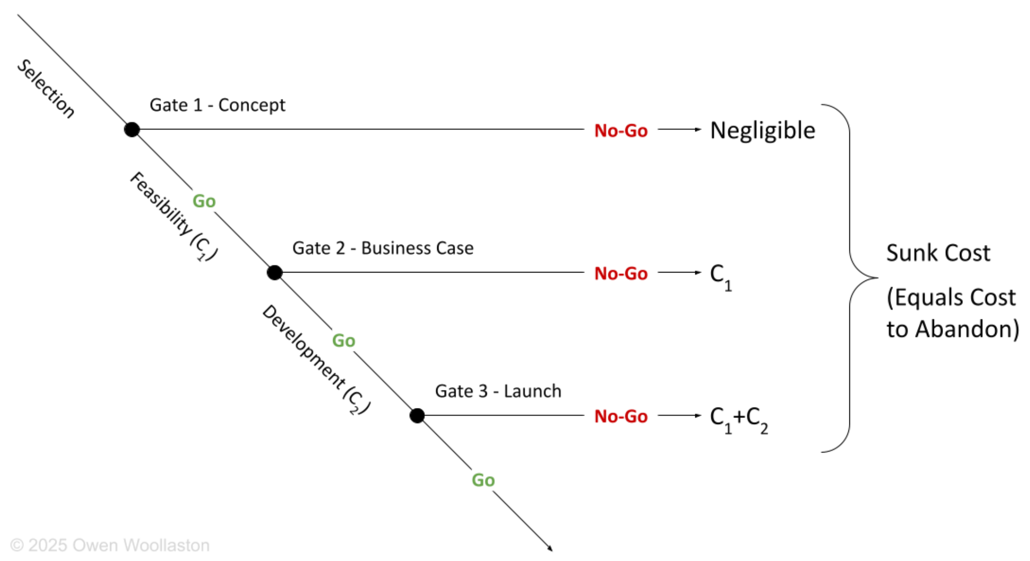


Figure 1 (an original work) depicts the process of transforming a product idea (beginning with a selection stage on the left) into a development project. This diagram positions R&D as a learning process in which uncertainty is translated into manageable risk and focuses on decision-making during a project (after selection). One gate has been magnified to highlight the (real) options typically available to R&D managers. At each gate, the state of the project is assessed, and options are identified as indicated under the magnifying glass in Figure 1. Generally, a product council may elect to *continue* the development as planned, *pivot* to a new technology or market, *resize* the project by adjusting the level of investment, *defer* (until certain conditions are met) the project, or *abandon* it altogether (Coldrick et al., 2005; Huchzermeier & Loch, 2001; Kauffman et al., 2015). These are the *real options* of R&D, and differ from stock options in their complexity and degrees of freedom.

Figure 2

Options Tree for R&D



Researchers of real options theory in R&D (Jagle, 1999; Neely III & De Neufville, 2001; Posen et al., 2018) often present a decision tree as an "options tree" similar to Figure 2, which includes a main success scenario and early exits where the project is abandoned. This relatively simple structure is unlikely to complicate communications, but it has hidden complexity. Firstly, the feasibility cost (C_1) and development cost (C_2) are rarely known ex-ante and must be progressively elaborated as the project proceeds. Secondly, the go/no-go decisions at all gates are ultimately based on expectation, both in the psychological and statistical senses of the word. Classical real options proponents would have managers calculate the expected return at each gate, discounting at each stage, to determine the project's expected value (Perlitz et al., 1999; Neely III & De Neufville, 2001), whereas contemporary proponents have begun incorporating behavioural elements to the mathematical approach (Posen et al., 2018).

This thinking leads to quantitative analysis, reducing the problem from its natural (social) domain to an apparently mathematical problem, requiring us to implicitly or explicitly evaluate the probability of failure at each gate. However, these probability evaluations remain primarily cognitive and social processes, subject to biases and heuristics as any other, so the appeal of analysis can backfire without adequate data supporting it. Similarly, Neely III and De Neufville (2001) warn that misuse of NPV and analytical methods introduces errors, and Kahneman and Tversky (1977) have amply demonstrated that human assessment of probability has several failings. Cost estimation in high-tech R&D is touched on in Part B; for now, it is worth noting that significant errors are common (McConnell, 2006).

With real options and other analytical approaches not addressing cognitive errors and other portfolio management tools being largely focused at the strategic level, the R&D portfolio management literature does not yet appear to address the application of decision-making psychology to R&D decision-making at a level deep enough to inform go/no-go decisions.

2.1.2.3 Commitment Escalation in R&D Decision-Making

Cognitive errors, especially biases, often lead to commitment escalation, where managers increase spending even as the project struggles or fails. Rather than withdrawing resources, management goes “all in” and “throws good money after bad” (Garland, 1990, p. 728). Commitment escalation is "... a common and pervasive phenomenon in NPD" (Sarangee et al., 2014, p. 2014).

Commitment escalation is a multifaceted phenomenon. A primary contributor is impression management – the pressure we experience to deliver on our commitments to be seen as consistent and trustworthy (Dorison et al., 2022; Cialdini, 1993) and thus maintain our reputation and goodwill. Bazerman (1994, p. 82) states, "We escalate because of our own previous commitments." Further, false beliefs and commitments can arise from false confidence early in a project or failure to effectively manage uncertainty and stakeholder expectations (Galasso & Simcoe, 2011). Prospect theory (Kahneman & Tversky, 1979) predicts loss aversion; that is, managers will perceive abandoning a project as more significant than resuming a different project of equivalent merit. Eliens et al. (2018) point to overconfidence: "A strong positive belief about one’s decision at the start of the project is necessary for escalation of commitment to occur."

Researchers sometimes differ on the significance of factors. Brockner (1992) identified that sunk cost is not a requirement for commitment escalation, yet Yang et al. (2020) observe, “[...] despite the negative feedback, the closer to the project completion and product launch, the more likely managers will invest further to the project to justify the sunk cost they already committed to the project and realise the potential opportunity they perceived". Roeth et al. (2019) conclude that a bad mood (negative affect) increases commitment escalation in some contexts but not others. While there is some clarity about significant factors – impression management, loss aversion, and overconfidence – the academic landscape remains somewhat confusing and contradictory about other factors, such as affect and sunk cost.

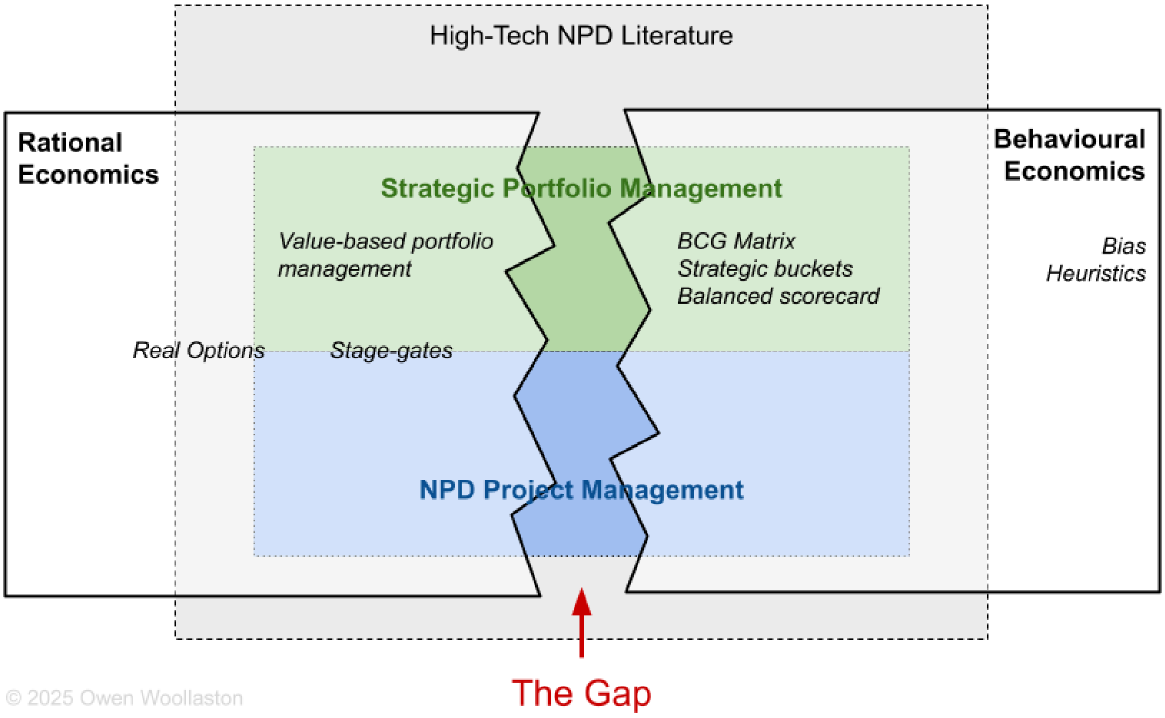
While the academic debate around factors continues, managers face another challenge: Commitment escalation is elusive in that it is difficult to identify, particularly during the project. As Eliens et al. (2018) noted, "[...] in both literature and practice, it often remains unclear when escalation of commitment is actually escalation or appropriately allocated commitment." Uncertainty clouds managers' ability to identify escalation in complex business environments. Worse, it may be impossible to determine retrospectively whether the investment was appropriate. Further clouding managers' ability to clearly identify commitment escalation is choice-support bias, in which post hoc rationalisation of decisions occurs (Lind et al., 2017). Haidt (2012) observed: "The mind is divided, like a rider on an elephant, and the rider's job is to serve the elephant," by which he means people tend to rationalise their decisions after the fact and not necessarily beforehand. Haidt has been paraphrased as saying: "We aren't rational; we're rationalising." This means that when it comes to commitment escalation, managers may not be aware they're doing it and may make up reasons why they did it.

There are a few approaches to mitigating commitment escalation. In their article "Attenuating the Escalation of Commitment to Faltering Projects in Decision-Making Groups: An Implementation Intention Approach," Wieber et al. (2015) advocate self-distancing by taking the perspective of an independent observer and managers pre-committing as a group to "implementation intentions" to making optimal and objective investment decision so that commitment escalation is mitigated by a pre-existing commitment to behave objectively. Eliens et al. (2018) advocate a similar strategy of gatekeeper independence in which decision-makers are divorced from the emotional investment in the project. The difficulty with these approaches is that they require invested managers to forego some decision-making power for the sake of objectivity. This type of decoupling is achievable in large organisations, but in high-tech SMEs, it might not be possible to get enough distance from the decision. This is because dual roles are common: The founder might also be the CTO or CEO. Alternatively, the board might have been involved with the portfolio investment decision in the first place. External consultants notwithstanding, there is often no one left in a position of power who doesn't also have a stake in the project.

In summary, commitment escalation is pervasive, multifaceted and elusive. A full treatment of the topic is beyond the scope of this thesis, but it is sufficient to highlight that it is a significant threat to effective go/no-go decision-making in R&D. The main problem it introduces in managers is the tendency to continue as planned and be adverse to abandoning projects. While mitigating approaches exist, managers' biases frequently undermine the objectivity required to align projects to strategic and portfolio goals, often because managers want to keep their project alive, appear competent and credible with stakeholders, or are overconfident in their decision-making abilities.

Figure 3

Research Gap



2.1.3 Summary of R&D Decision-Making Literature and Research Gap

As outlined above, several portfolio management approaches exist to facilitate portfolio investment decision-making (in italics in Figure 3). Many operate at a portfolio strategy level, providing only guidance on strategic alignment to go/no-go decision-makers. Of the portfolio-management frameworks identified, only real options theory provides specific insight into mid-project decisions, the available options, and how to evaluate them. However, real options theory is a quantitative model and does not provide guidance on mitigating commitment escalation or a structured approach to go/no-go decision-making. Figure 3 illustrates the gap in two dimensions. Vertically, stage-gate models and real options theory connect the project and portfolio management disciplines by directly focusing on mid-project decision points. However, there is little overlap in the high-tech R&D literature horizontally between rational economics on the left (including real-options and value-based portfolio management) and behavioural economics on the right (including the psychology of decision-making, commitment escalation, bias and heuristics).

Thus (1) a systematic review of the R&D literature is required to clarify and validate this gap, and (2) a framework for go/no-go decision-making will contribute to the literature by integrating knowledge distributed across multiple disciplines. This approach will directly address the first two aims of this thesis, restated here:

1. **Conduct an AI-assisted literature review** to identify existing theories related to go/no-go decision-making in high-tech R&D.
2. **Develop an integrated decision-making framework** based on insights from the literature, providing structured guidance for go/no-go decisions.

Without an integrated decision-making framework, managers must rely on disparate tools and theories that do not fully address the interplay of portfolio strategy, risk evaluation, and commitment escalation. This thesis seeks to fill that gap.

2.1.4 Theoretical Constructs

Figure 1 above illustrates "Gates and Options in R&D". This diagram provides useful context but is not a sufficient basis on which to perform a literature review. For this, it is necessary to identify the constructs that will underpin the review.

Drawing from the theoretical foundations above, this thesis focuses on five key constructs:

Antecedents are things in place that precede the decision-making process and decision point. Antecedents are present in several forms, an example being capabilities. Examples of capabilities include organisational agility (Vaculik et al., 2019) and dynamic capabilities (Teece et al., 1997; Teirlinck & Spithoven, 2012). These capabilities, and antecedents in general, are characterised by the need to develop them a priori. Stopping to do squats in the middle of a marathon is counterproductive. A runner must achieve fitness before the event, just as a firm must have the necessary strategy and capabilities before they are called on.

Approach is a multifaceted concept that refers to the decision-making methodology – *how* the decision is made – and may be influenced by the decision-maker's philosophy, beliefs, past experiences and personal preference. Approaches can be process or tool-based, formal or informal, group or individual, analytic or intuitive. Some examples of approaches include hybrid methods such as multi-criteria decision-making (Mardani et al., 2015) and quality function deployment (Rianmora & Werawatganon, 2021).

Criteria are specific factors to consider when weighing a decision. Criteria can be qualitative or quantitative, and common criteria in R&D decisions include development cost, expected return, level of risk, and strategic alignment.

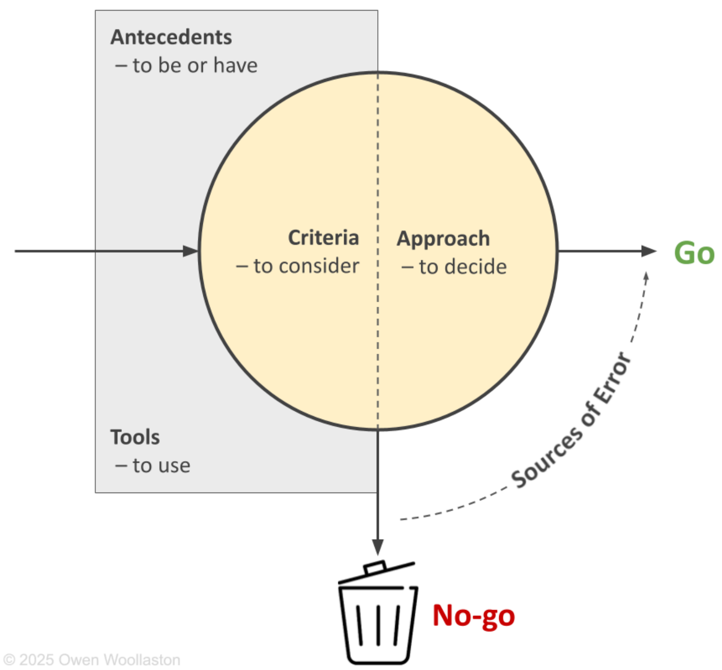
Sources of Error are factors that could (knowingly or unknowingly) introduce error (usually bias), as discussed in the section on commitment escalation above.

Tools are mental or physical devices that help improve decision-making quality, particularly in measuring, preparing and interpreting information critical to the decision. Examples of tools in R&D include sensitivity analysis (Smit & Trigeorgis, 2007), simulation (Cooper, 2006; Cooper et al., 2000), and decision trees (Jagle, 1999; Neely III & De Neufville, 2001). Tools can help quantify, conceptualise and clarify options and criteria.

Options are a concept that would normally feature prominently in a thesis about decision-making. However, this thesis has only two options: go or no-go. In this context, options are not considered a theoretical construct because there are no degrees of freedom. In a more relaxed sense, the reader may still consider "go" to mean continue, pivot or resize in the sense of real options theory as applied to R&D by Smit and Trigeorgis (2007). Similarly, "no-go" can be taken to mean defer or abandon.

Figure 4

Skeletal Framework Based on Theoretical Constructs



These five theoretical constructs – antecedents, approach, criteria, tools, and sources of error – provide a taxonomy to classify factors influencing R&D decision-making, forming the basis for the skeletal framework shown in Figure 4. This framework positions antecedents and tools as preceding the decision point, with criteria and approach central to the decision itself, while sources of error are depicted on a dotted arc from no-go to go to indicate managers' tendency toward commitment escalation. These constructs will also guide the synthesis of the literature. Combined with the theoretical foundation, the skeletal framework establishes the groundwork for the literature review and the development of a go/no-go framework for R&D decision-making. However, before this synthesis can proceed, a literature review method, including the AI-assisted approach, must first be developed.

2.2 Literature Review Approach

Having established the theoretical foundations, it is useful to reexamine the aims of the thesis to clarify the literature review's scope and purpose. The first aim was to: "Conduct an AI-assisted literature review to identify existing theories related to go/no-go decision-making in high-tech R&D." This aim specifically focused the literature review on go/no-go decision-making and limited the search domain to R&D in the technology sector. Further, it focuses the research method on using AI-assisted literature search.

Once the literature search was completed, the results were synthesised per the second aim: "Develop an integrated decision-making framework based on insights from the literature, providing structured guidance for go/no-go decisions" – this is covered in the synthesis methodology section. Finally, the discussion and conclusion highlight other potentially relevant theories not identified in the literature search.

The systematic approaches summarised in Saunders et al. (2022) and in more detail by Xiao and Watson (2019) were followed throughout this section.

2.2.1 Research Philosophy

The philosophy behind the review was pragmatism. The search aimed to identify factors relevant to go/no-go decision-making from any field of study that illuminated go/no-go decision-making, using a novel AI-driven approach to expand the scale of the search. Casting a wide net to find useful insights is the defining characteristic of pragmatism (Saunders et al., 2022; Sun & Zuo, 2024; Mir & Greenwood, 2021), which, in this thesis, was driven by the intention to develop a practical, helpful framework for practitioners.

At times, the review borrowed a different philosophical stance. For example, by recognising biases and heuristics that drive commitment escalation in the theoretical foundations, this thesis also acknowledged that the way managers perceive the state of their projects can differ from objective reality. Acknowledging different perspectives is called *critical realism* (Saunders et al., 2022; Sun & Zuo, 2024; Mir & Greenwood, 2021). However, in this case, critical realism was not the underlying philosophy. It was simply a useful way to look at biases, and this temporary adoption of critical realism was just another expression of pragmatism. Thus, the literature review took a pragmatic stance to develop a practical method to answer a practical research question that would benefit practitioners.

2.2.2 Traditional Searches and Limitations

Established methods of literature search follow a process of selecting publications, refining keywords and iterating until the number of search results fits the research scope and the research team's capacity. Right-sizing the review is necessary to complete the research within its budget and schedule, whether dictated by grant funding or the number of months and points allocated to the degree programme. Online metadata services (e.g. Scopus) are then used to filter millions of articles by keyword, publication date, publication type, citation count or author, the objective being to reduce the size of the search results to a relevant yet manageable list (Saunders et al., 2022).

The problem with this method of searching is that it takes time and effort to filter through irrelevant articles to identify the relevant ones. In the traditional approach, keyword matches (which may be exact or use wildcards) and boolean logic (AND, OR) are typically used to filter articles by matching against title, author or abstract. These are *syntactic matches* where the characters in the keywords must match the characters in the article, and the search engine is robotic in that it has no concept of meaning or simile. *Semantic matches* – where the *meaning* of the search term is matched – do not work, so human researchers must reject numerous articles to locate relevant articles. Researchers can reduce the number of irrelevant articles by using highly selective search parameters, but this increases the chance of omitting relevant articles. As the body of literature has increased, a catch-22 situation has emerged where researchers must choose between speed and completeness – either unknowingly rejecting relevant articles or working through many irrelevant search results (Joos et al., 2024; Mourão et al., 2020). The confusion matrix in Table 1 summarises the errors when searching the literature.

Table 1

Confusion Matrix for Traditional Keyword Searches

	Excluded in Search	Included in Search
Irrelevant to Research	True negative (TN) – irrelevant articles excluded from the search.	False positive (FP) – irrelevant results are noise that takes time and effort to remove manually.
Relevant to Research	False negative (FN) – an unknown number of relevant articles excluded from the search.	True positive (TP) – relevant articles included in the search.

Given this catch-22 that researchers increasingly experience, there is a growing need to be more efficient and discerning when searching the literature. This study explores AI-based approaches as a possible solution.

2.2.3 Introduction to AI-Assisted Search

Traditionally, the way to reduce irrelevant articles in search results was to be more selective and thus reduce the sample size, but this increased the chance of rejecting relevant articles, and vice versa. In this thesis, an AI-assisted approach was expected to reduce false negative (FN) and false positive (FP) errors by implementing a more discerning filter based on the article's AI-reported relevance to the research question. Higher discernment was expected to lead to better discrimination of candidate articles and fewer false positives. Furthermore, the automated approach was expected to process thousands of articles quickly, making narrowly defined keyword searches unnecessary. False negatives were expected to be reduced by increasing the search space. These expected improvements are captured qualitatively in Table 2 and verified empirically in Appendix G.

Table 2

Expected Confusion Matrix for AI-Assisted Search

	Excluded in Search	Change due to AI	Included in Search
Irrelevant to Research	Increase in true negative results		Reduction in false positives
Relevant to Research	Reduction in false negatives	Increase in true positive results	

The argument for exploring the use of AI was compelling, and the timing for this thesis – started in 2024 – was perfect. With the general availability of ChatGPT in November 2022, which offered an interactive chat dialogue interface to OpenAI’s large language model (LLM) GPT-3.5, many researchers began exploring new methods and supporting research, particularly literature reviews (Tóth & Oldal, 2022; Wagner et al., 2022; Joos et al., 2024). This chat interface allowed researchers to upload articles and request interpretations and summaries interactively, accelerating learning by relying on AI to digest and reinterpret voluminous information.

While useful for learning, the chat interface had limitations for large-scale support of systematic literature reviews. First, scale was an issue – only ten documents could be uploaded for review at any time. Second, there were ethical and legal complications when using the output. Sometimes, the AI responses quoted word-for-word, unknowingly breaching copyright, and at other times, the output was inaccurate or omitted important facts and figures (Qureshi et al., 2023). The provenance of the information was sometimes questionable when generated by AI, and most universities viewed it as unethical to present AI-generated material as original work. Third, process integration was poor when interacting with a chat interface, requiring researchers to interact to get responses and limiting automation. Fourth, citing AI-generated work was messy and required copying the output into an appendix because – fifth – the repeatability was poor: There was no guarantee that the AI would generate the same output a second time. Sixth, and finally, for the aforementioned reasons, the academic credibility of chat responses was low. (Qureshi et al., 2023; Yao et al., 2024).

This thesis sought to explore the use of AI to accelerate literature reviews without compromising the work's originality, accuracy, repeatability or credibility. To achieve this, the method was focused on developing a programmatic, semi-automated approach to accelerating the search process. This semi-automated approach was preferable to traditional methods or using AI chat interfaces because it implemented a more discerning filter than traditional keyword searches (based on conceptual relevance), enabled the search to scale up without increasing the burden on researchers, and, once configured, required only a modicum of human supervision.

As the effectiveness of the overall approach hinged on the quality and rigour of the search, the following section focuses on detailing this component.

Table 3

Balancing Ethics and Efficiency in an AI-Assisted Review

Phase	Actor	Approach	Rationale
Planning	This Author	Test the search space and develop	Some testing of the search parameters is still necessary.
Search	AI with this author's supervision	Use broad metadata searches to leverage automation and avoid unnecessarily restricting the search space. Fetch metadata automatically based on search terms, then rank and filter based on AI-generated relevance to the research question.	Automation reduces manual effort and enables a bigger search space. The research question better indicates relevance than keywords or other metadata.
Selection	This Author	Manually rate and select the remaining articles based on AI relevance.	This author's review of search results is required to meet ethical requirements.
Data extraction Synthesis Quality assessment Reporting	This Author	Manually use established methods	Avoid the ethical grey zone around the originality of the thinking and output.

2.2.4 Search Approach and Ethical Considerations

While AI-assisted methods offer promising improvements in the scale and efficiency of searches, they also raise ethical considerations that must be addressed to ensure academic integrity. To fully understand the ethical considerations, it was necessary to specify where and how AI would be used. To do this, the AI approach was reviewed against the phases of a systematic literature review as described by Xiao and Watson (2019) – see Table 3 above. The proposed method began with traditional search planning. Then, as indicated by the green highlighted section, AI was used as a discerning filter, followed by traditional methods of article selection, data extraction, synthesis, and reporting.

Once specified in the context of the review process, the University of Waikato's "Guidelines for Student Use of Generative AI Tools" was used to verify the ethical approach (University of Waikato, n.d.). The guidelines require students to:

1. **Critically evaluate AI outputs** – naturally achieved as part of the process because the AI-search step flowed into the manual selection of articles.
2. **Use AI in a supplementary way** – inherently achieved because AI was not used as the primary source of information – it was only used to identify primary sources.
3. **Adhere to academic integrity.** This approach "played it safe" by using AI solely as an improved search tool.

As an aside, this author envisaged that other steps in the review process were technically possible to implement using AI (such as data extraction and synthesis). However, this thesis avoided the ethical grey zone introduced by having AI do any of the "thinking" because this would have eroded any claim to originality and introduced ethical complications.

2.3 AI-Assisted Literature Review Methodology

This section expands on the methodology's details, continuing to follow the guidance of Xiao and Watson (2019), including the specifics of the search method, the selection of an academic metadata service, search parameters, and search history. It also describes the AI-assisted search process and its interactions with OpenAI's application programming interface (API). Next, it describes the process of identifying and screening articles, followed by defining data items. The section concludes by describing the synthesis methodology.

2.3.1 Metadata Service Selection

The goal was to integrate AI with academic metadata services to produce an AI-assisted search engine. This approach was chosen because time savings (through automation) were anticipated, and the system could scale to review thousands of papers. However, every academic metadata service had a different API, and writing the software to integrate with APIs was time-consuming. Therefore, the decision was made to integrate only one metadata database. Wikipedia (2024) publishes a list of metadata services and full-text academic databases, which was used as the starting point for source selection (Mesgari et al., 2015). These services are summarised in Table 4.

Table 4

Metadata Services - from Wikipedia (2024)

Name	Size	Providers	API Access
Google Scholar	389,000,000	Google	None
BASE: Bielefeld Academic Search Engine	259,497,239	Bielefeld University	Access restrictions.
OpenAIRE Graph	242,000,000	OpenAIRE AMKE	Free, but open pub only
CORE	207,000,000	Open University	Free - rate limits
OpenAlex	205,000,000	OurResearch	Free
Crossref	111,231,936	Crossref	Free
MyScienceWork	90,000,000	MyScienceWork Inc	Free
Scopus	78,000,000	Elsevier	Subscription
ScienceOpen	66,000,000	ScienceOpen Inc.	Free
National Diet Library Collection	44,187,016	National Diet Library	Free
Russian Science Citation Index	35,409,829	Scientific Electronic Library	Free
OAIster	30,000,000	OCLC	Free

Because the intention was to develop an AI-assisted search tool, services needed to meet specific requirements:

1. Provide article metadata, including title, abstract, and publication date.
2. Be accessible via the University of Waikato.
3. Publish an application programming interface (API).
4. Offer extensive collections of relevant publications.
5. Be multidisciplinary, broadly including economics, psychology and management.

Scopus was chosen because it met all the above criteria and was available via the University of Waikato. Others that included free API access were avoided because free services often implement rate-limiting – they limit the number of queries – and Google Scholar did not implement an API, making it unsuitable for programmatic access.

With the metadata source determined, the next step was to define the search parameters to ensure the retrieval of relevant and high-quality literature.

2.3.2 Search Parameters

Having identified Scopus as the metadata service, the next step was to define the search parameters and the services they used to implement the search (see Table 5). This design defined journals, keywords, minimum citation count and earliest year of publication in the same way as online Scopus searches. The final two parameters – review keywords and research question – were used to instruct the AI to calculate the relevance of each article. The details of the AI-based ranking are in the next section. These six parameters provided the complete specifications for the literature search.

Table 5

Search Design Parameters

Parameter	Description	Value	Service
Search journals	List of publications to include in Scopus query	Fixed list used for all searches See Appendix A	Scopus
Search keywords	Boolean (AND/OR) combination of keywords passed in Scopus query	Refer to the history of Scopus searches in Table 6 below	Scopus
Earliest year	Include results from the earliest year up until the present day.	Refer to the history of Scopus searches in Table 6 below	Scopus
Citation percentile	Includes articles predating the earliest year if they are cited above this threshold.	95% (for all searches)	Scopus
AI review keywords	Simplified list of keywords without boolean operators suitable for passing to AI for natural language processing	decision-making, R&D, innovation management, product management, new product development, high-tech, escalation of commitment (for all GPT-4o reviews)	OpenAI
AI research question	The research question for AI to use to assess article relevance	How should firms evaluate whether to continue or abandon projects in high-tech R&D? (for all GPT-4o reviews)	OpenAI

Table 6

Final Scopus Queries

#	Time of Search	Scopus Query[2]	Year	Scopus	New
1	14/04/2024 15:32:21	(TITLE-ABS-KEY(decision making) AND TITLE-ABS-KEY(research and development) OR TITLE-ABS-KEY(R&D) OR TITLE-ABS-KEY(product portfolio management) OR TITLE-ABS-KEY(innovation management) OR TITLE-ABS-KEY(product management) OR TITLE-ABS-KEY(new product development) AND TITLE-ABS-KEY(technology) OR TITLE-ABS-KEY(high-tech)) AND <JOURNAL LIST>))	2010	776	776
2	21/04/2024 10:06:38	((TITLE-ABS-KEY(systematic literature review) AND TITLE-ABS-KEY(innovation strategy))) AND <JOURNAL LIST>))	2020	81	80
3	21/04/2024 13:58:33	(((TITLE-ABS-KEY(innovation strategy) OR TITLE-ABS-KEY(product management)) AND TITLE-ABS-KEY(abandon))) AND <JOURNAL LIST>))	1990	17	16
4	21/04/2024 15:57:33	DOI(10.1109/TEM.2018.2798922)))	1990	1	1
5	21/04/2024 15:58:16	DOI(10.1177/002224379703400114)))	1990	1	1
6	02/06/2024 12:21:42	((TITLE-ABS-KEY(decision making) AND (TITLE-ABS-KEY(research and development) OR TITLE-ABS-KEY(R&D) OR TITLE-ABS-KEY(product portfolio management) OR TITLE-ABS-KEY(innovation management) OR TITLE-ABS-KEY(product management) OR TITLE-ABS-KEY(new product development)) AND (TITLE-ABS-KEY(technology) OR TITLE-ABS-KEY(high-tech)))) AND <JOURNAL LIST>))	2000	1183	407
7	02/06/2024 12:37:03	(((TITLE-ABS-KEY(decision making) OR TITLE-ABS-KEY(project management)) AND (TITLE-ABS-KEY(research and development) OR TITLE-ABS-KEY(R&D) OR TITLE-ABS-KEY(product portfolio management) OR TITLE-ABS-KEY(innovation management) OR TITLE-ABS-KEY(product management) OR TITLE-ABS-KEY(new product development)) AND (TITLE-ABS-KEY(technology) OR TITLE-ABS-KEY(high-tech)))) AND <JOURNAL LIST>))	1900	2775	1586
8	14/06/2024 20:45:18	(((TITLE-ABS-KEY(escalation of commitment)) AND (TITLE-ABS-KEY(new product development) OR TITLE-ABS-KEY(NPD)))) AND <JOURNAL LIST>))	1900	14	10
TOTAL				4848	2877

[2] Appendix A defines the <JOURNAL LIST>.

2.3.3 Article Identification

Designing the AI-assisted search system was a non-trivial activity that required many iterations of the design and software. The search strategy evolved iteratively to refine keyword selection and ensure a comprehensive dataset. Table 6 above presents the final refined queries that yielded the most relevant results, following adjustments based on exploratory searches. This re-run meant that two series of searches were captured, the first of which (the exploratory series) is documented in Appendix B. The column “Scopus” indicates the number of results returned in the Scopus query. The column “New” shows the number of articles added after duplicates were automatically eliminated. “Year” indicates the year of publication. In total, 4848 records were fetched from Scopus (including two from snowballing), resulting in 2877 unique records in the local database.

Note the absence of PUBYEAR in the Scopus queries. This was not an error. Filtering based on the citation percentile was only possible if all years were first downloaded from Scopus and then filtered in the software. For example, in Table 6 above, the number of items downloaded from Scopus for row seven was 2775. However, after filtering by the earliest year and highest cited and eliminating duplicates, this was reduced to 1586.

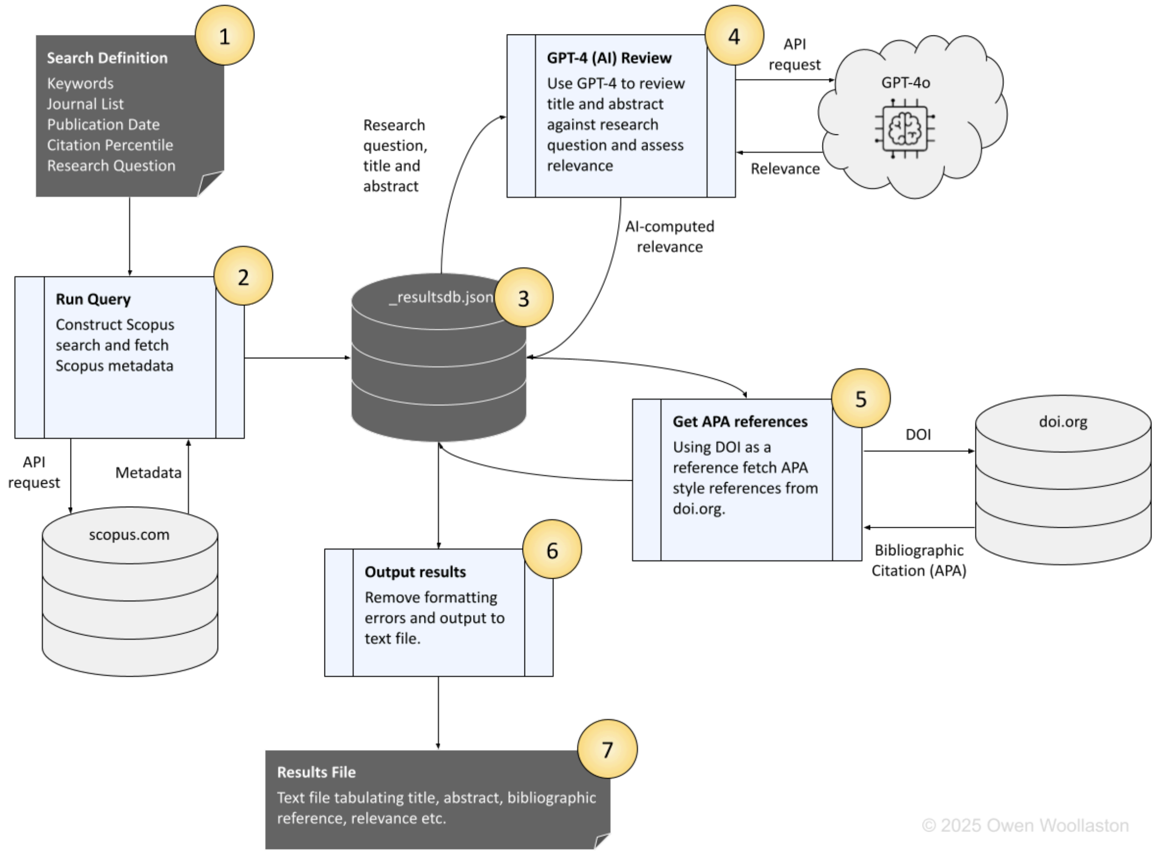
2.3.4 AI-Assisted Search System

The AI-based portion of the search was implemented using OpenAI’s chat API via the Python Module OpenAI version 1.17.0 (OpenAI, 2024). The specific model (version) used was GPT-4o. This model was publicly available and cost-effective when the search was run, although initial design work had begun on an earlier model.

This author implemented a semi-automated, programmatic interface to facilitate bulk searching and pre-screening of articles from Scopus metadata. From the researcher’s perspective, to start the search process, a configuration file was edited to set the search parameters from Table 5 above. Then, Scopus was queried to test and download each article’s metadata. Next, the script augmented the metadata with the AI-computed relevance to the research question (*AI_R*). Finally, the researcher downloaded the results for manual sorting and filtering by relevance and other criteria, and the traditional literature review steps continued. The mechanics of this process are detailed in Figure 5 below.

Figure 5

AI-Assisted Search Process



The implementation consisted of a set of Python (Python Software Foundation, 2024) scripts in a GitHub repository that were run on the command line inside a GitHub Codespace (GitHub, n.d.). Each script performed a specific action (usually interacting with a web service) and updated a database. The scripts did not need to be run in sequence, but there was a sensible order in which to run them, as outlined in Figure 5:

1. The researcher edited a file named `search_design.py` to define the search parameters in Table 5 above.
2. Scopus was searched by running the script `scopus_query.py --download`. Running it without `--download` reported the query size but did not download the results, which saved time when performing initial keyword exploration.
3. The new result set was added to the `_resultsdb.json` database, and duplicates were discarded. The details of the Scopus query were appended to a file named `_query_history.txt`, which tracked the history of searches.

4. The researcher invoked `gpt_review.py` to use GPT-4o to review new articles. The database `_resultsdb.json` was updated to include each article's AI-calculated relevance.
5. The researcher invoked `doi_query.py`, which fetched correctly formatted APA references from https://doi.org and added them to `_resultsdb.json`. Reference lookup was a nice-to-have functionality unrelated to the AI part of the process.
6. The results were processed by `save_results.py`, which generated the file `results.txt` in tab-delimited format suitable for copying into a spreadsheet.
7. The researcher resumed reviewing the AI-assisted list of results using a spreadsheet or other reference management tool.

The end-to-end process could have been fully automated so that the researcher-defined `search_design.py` and a single script generated the file `results.txt`; however, this risked incurring unnecessary costs using OpenAI's services and increasing the chance of errors because the scripts had not been production-hardened and, therefore, required a degree of handholding to run successfully.

2.3.5 AI Screening and Relevance Scoring

Following the implementation of the software process and scripts, it was necessary to define the specific instructions required to instruct the AI to rank the relevance of an article. The GPT-4o model, accessed via the OpenAI Python module, required two arguments:

1. `system_command` – instructions to the AI model to define its behaviour
2. `user_command` – the content or question to be processed

After some experimentation, it was found that the system_command below was effective:

```
"My research question is '<research_question>' and my keywords are <keyword_list>.  How relevant is this abstract to my research? Respond using only a 1-10 scale."
```

The user_command followed this format:

```
"Title: <title> Abstract: <abstract>"
```

When provided with these instructions and content, GPT-4o responded on a 1–10 scale, with 10 being the most relevant. Qureshi et al. (2023), Burger et al. (2023), and Jin et al. (2024) all identified similar methods used by other researchers, in which ChatGPT was used to assess relevance to the research question.

The AI ranking process was applied to the 2877 articles identified through the initial automated Scopus searches. With a cutoff relevance of $AI\_R \geq 7.0$, the process returned 81 articles. This cutoff was found by trial and error to strike an optimal balance between the number of articles to review (researcher's effort) and the relevance of returned articles (selection rate). A higher cutoff returned too few articles, and a lower cutoff produced too much noise and increased the review effort without significantly increasing the number of relevant articles.

Table 7

Article Eligibility Criteria

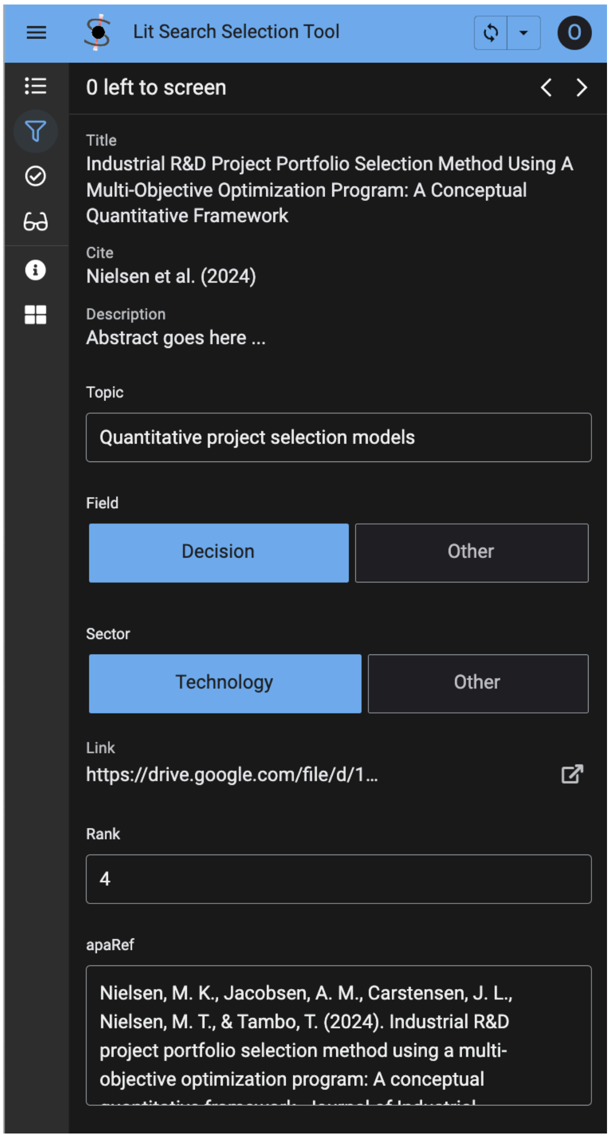
	Operator	Characteristic	Constraint	Rater
Inclusion Criteria	Logical AND	Relevance (*AI_R*)	≥ 7.0	AI
		Language	English	Human
		Availability	Full-text	Human
		Academic field	Decision-making or related	Human
		Industry sector	Technology or non-specific	Human
		Rank (relevance)	≥ 4	Human
	Logical OR	Year of publication	Any (relied on *AI_R* and rank)	
		Citation percentile	Any (relied on *AI_R* and rank)	
Exclusion Criteria (logical negation of inclusion criteria)	Logical OR	Relevance (*AI_R*)	< 7.0	AI
		Language	Non-English	Human
		Availability	Abstract-only or Unavailable	Human
		Academic field	Unrelated to decision-making	Human
		Industry sector	Industry-specific and not Technology	Human
		Rank (relevance)	< 4	Human

2.3.6 Manual Screening and Eligibility

With 81 articles having been screened for eligibility by the AI-based relevance, the process reverted to traditional eligibility review by this author using the criteria defined in Table 7. A tool was developed using Google AppSheet (AppSheet, n.d.) to assist in classifying articles, as pictured in Figure 6. The tool displayed the title and abstract and linked to the downloaded article, allowing this author to classify and rank the articles in situ and correct formatting errors in the automatically generated APA reference.

Figure 6

Article Classification Tool



All 81 articles were reviewed and classified according to the academic field and industry sector. Articles not relevant to decision-making in the technology sector were rejected, as were articles not in English or not available in full text. Then, the title and abstract were reviewed, and the body was skim-read to rank each article on a 1–5 Likert scale defined in Table 8. Articles ranked less than four were also rejected. Ranking resulted in 28 articles included in the review.

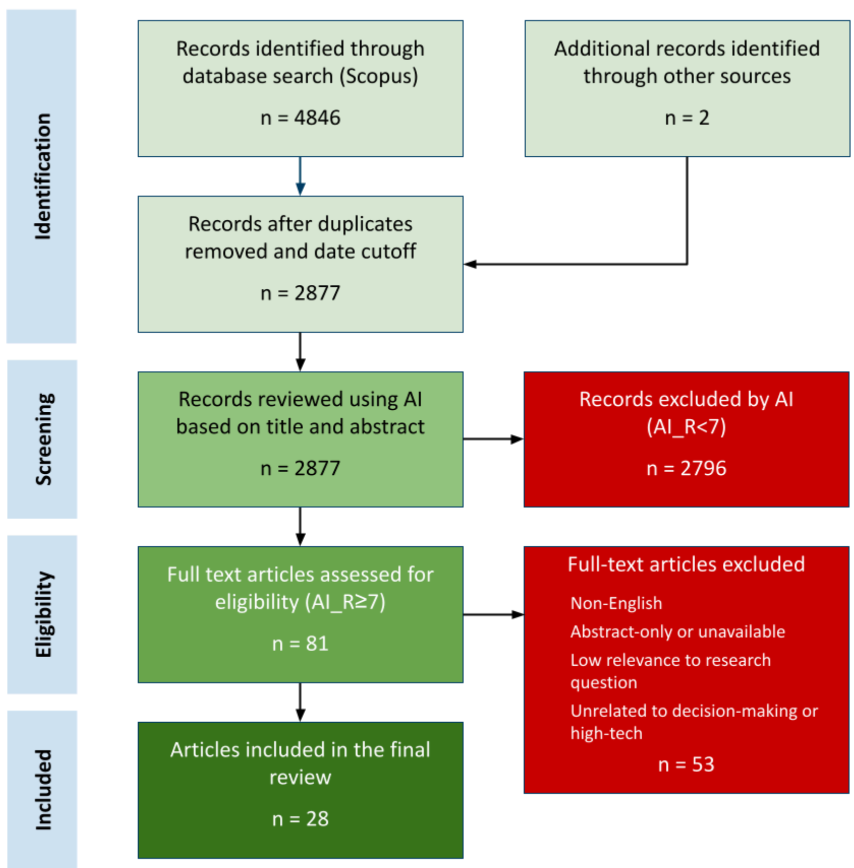
Table 8

Manual Ranking Scale

Rank	Definition
1	Irrelevant – completely wrong field
2	Irrelevant – right field, wrong topic
3	Irrelevant, but adjacent to the topic or near miss
4	Relevant – include in the review
5	Highly relevant – include in the review

Figure 7

PRISMA Diagram



The PRISMA diagram (Page et al., 2021) in Figure 7 summarises the complete search process. Recapping, 4848 records were identified on Scopus and other sources, of which 2877 were unique records. The AI-search method eliminated 2796 records, leaving 81 articles identified for eligibility and human rater screening, after which 28 articles were selected as relevant to the research question, in the technology sector and the field of decision-making. With these 28 articles, and the first aim (the literature review) completed, the thesis proceeded to extraction of data items (decision-making factors) and to synthesis of the framework, which is described below.

2.4 Synthesis Methodology

The second aim of this thesis was: Develop an integrated decision-making framework based on insights from the literature, providing structured guidance for go/no-go decisions. This is the primary goal of the synthesis method, described here.

With 28 selected articles, the thesis turned to extraction and categorisation of decision-making factors, followed by restructuring and grouping and finally synthesis of the framework. The synthesis process involved three broad stages: (1) extracting decision-making factors from the selected 28 articles, (2) categorising them based on theoretical constructs, and (3) structuring them into a conceptual framework. The goal was to systematically identify patterns and insights that inform R&D investment decisions.

2.4.1 Data Categorisation and Extraction

The 28 articles were reviewed carefully, categorising data and extracting key data points that would inform the synthesis of findings. The constructs from the theoretical foundations (summarised below) were adopted as categories against which to classify decision-making factors.

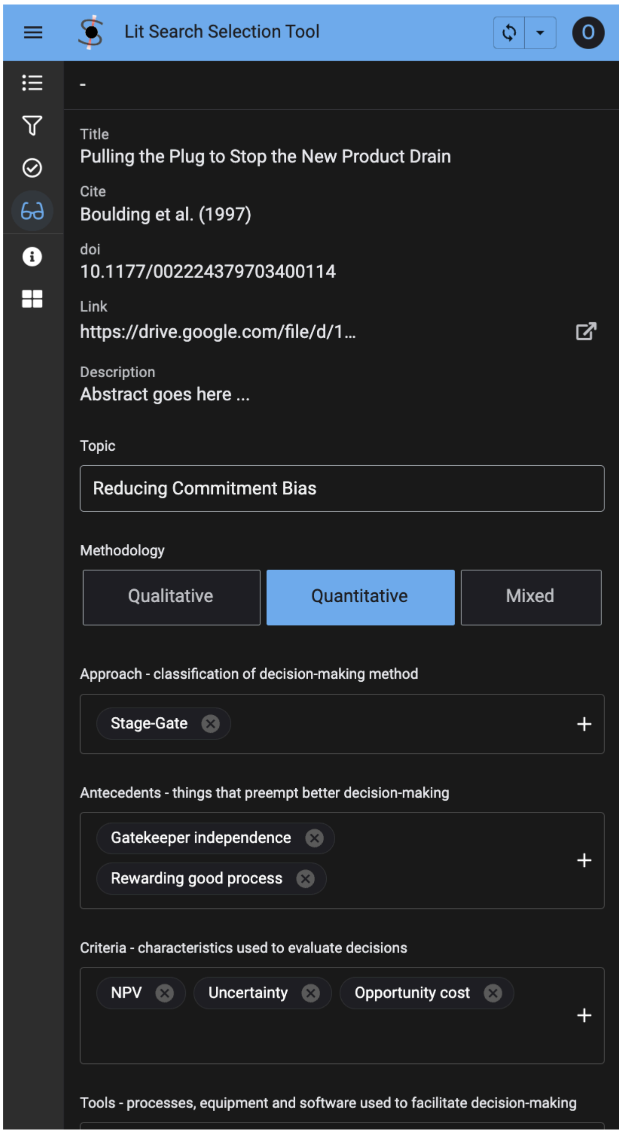
The theoretical constructs that became the data categories were:

1. Antecedents: Preconditions for decision-making, such as strategy and capabilities that must be established before investment decisions.
2. Approach: The methodology used in decision-making: shaped by philosophy, experience, and preferences and can be formal/informal, analytic/intuitive.
3. Criteria: The qualitative or quantitative factors considered when evaluating a decision, such as cost, risk, return, and strategic alignment.
4. Sources of Error: Factors that introduce bias or inaccuracies into the decision-making process, knowingly or unknowingly.

5. Tools: Mental or physical aids that enhance decision-making by measuring, analysing, and interpreting critical information.

Figure 8

Data Extraction Instrument



Then, a data extraction instrument (Figure 8) was used to facilitate viewing each selected article in turn and extracting data that fit the construct.

The data extraction instrument – developed using Google AppSheet (AppSheet, n.d.) – facilitated extracting each selected article's research methodology, approach, antecedents, criteria, tools and identified sources of error. The tool presented each article for review, including its title, abstract, full text, authors, publication date, rank, and AI-relevance score. It provided multi-select drop-down fields for each category – approach, antecedents, criteria, tools, and sources of error – allowing zero, one, or multiple values to be entered against each category, with the option to add new values if needed. For example, this author could enter “rational” in the decision-making approach field for an article and add “real options” if appropriate to the article's content. In this way, a database of codified characteristics was constructed for the selected articles.

2.4.2 Structuring Extracted Data

Once the data extraction process was completed, a data table conforming to the schema in Table 9 was produced. By “schema”, it is meant that each row in Table 9 defines a column of data in the results table, and each row in the results represents a selected article. No extra work was required for this step because, by design, the data extraction instrument natively stored the results in this format.

Table 9

Data Schema

Field	Data Format	Example Value
Cite	Text – in-text citation referring to article	Bloggs et al. (1997)
DOI	Text – valid DOI number; used as primary key	10.1234/123123123
Methodology	Text – one of: Quantitative|Qualitative|Mixed	Quantitative
Approach*	Text – comma-separated list of values	Portfolio management, MOOP
Antecedents*	Text – comma-separated list of values	Innovation strategy, Market awareness, Risk orientation, Budget allocation
Criteria*	Text – comma-separated list of values	Benefits, Project costs, Strategic alignment, Balanced portfolio
Tools*	Text – comma-separated list of values	Scoring model, Graphical cost-scope-time
Error sources*	Text – comma-separated list of values	Not invented here, Risk aversion
PubYear	Number – the year of publication	2001

* data items of primary significance

2.4.3 Classification and Thematic Grouping

Each record (originally grouped by article) was then pivoted using a function developed in Google Apps Script (Google, 2024). This function splits the data items stored in the field, then groups the data items by value and collates the authors and summary statistics. For example, suppose a single truncated record was represented by Table 10a below. In that case, the contribution to the results was to add (or update) Portfolio Management and MOOP as approaches and to include Bloggs et al. (1997) in the list of authors (Table 10b).

Table 10a

Thematic Grouping – Original

Citation	Approach*
Bloggs et al. (1997)	Portfolio management, MOOP
(Other records omitted for brevity)	

➤

Table 10b

Thematic Grouping – Transformed

Approach	Citation/s	Count
+ Portfolio management	+ Bloggs et al. (1997), …	+ 1
+ MOOP	+ Bloggs et al. (1997), …	+ 1
(Other approaches from other records)		

A truncated example of the output generated is shown in Table 11, which lists in each row a decision-making criterion (e.g. NPV) and a citation for each article that refers to the item as a criterion. In this case, the table applies only to decision-making criteria. A separate table was generated for each category – antecedents, approach, criteria, sources of error and tools.

Table 11

Decision-Making Criteria

Criterion	Citation/s	Cites
NPV	Boulding et al. (1997), Jagle (1999), MacMillan & McGrath (2002), Nielsen et al. (2024), Perlitz et al. (1999), Pillai et al. (2002), Vaculik et al. (2019)	7
Benefits	Cooper (2006), Hauser & Zettelmeyer (1997), Jagle (1999), MacMillan & McGrath (2002), Neely III & De Neufville (2001), Nielsen et al. (2024), Perlitz et al. (1999)	7
…	…	…

A common issue was the large number of values for each data item and the different terminology used by researchers. For example, in the criteria category there were 35 values: Ambiguity, assumption validity, balanced portfolio, benefits, budget, capability, capacity, comparable asset valuation, competition, competitive advantage, constraints, costs, critical threat, customer benefits, disruptive rank, economic capital value, management support, market value, MOOP, NPV, opportunity cost, pace of technological change, performance threat, probability of commercial success, probability of technical success, project interdependency, project ranking, risks, shareholder value, strategic alignment, sufficient resources, supply chain, technical feasibility, transaction cost, and uncertainty.

This proliferation of values necessitated classifying data items into broader categories. The categories for criteria were confidence, constraints, culture, error sources, external pressure, finance, outcomes, portfolio, resource, strategy, and threats. Coarser classification, in turn, enabled tabulating and visually displaying results in a way that was easier to manage and present.

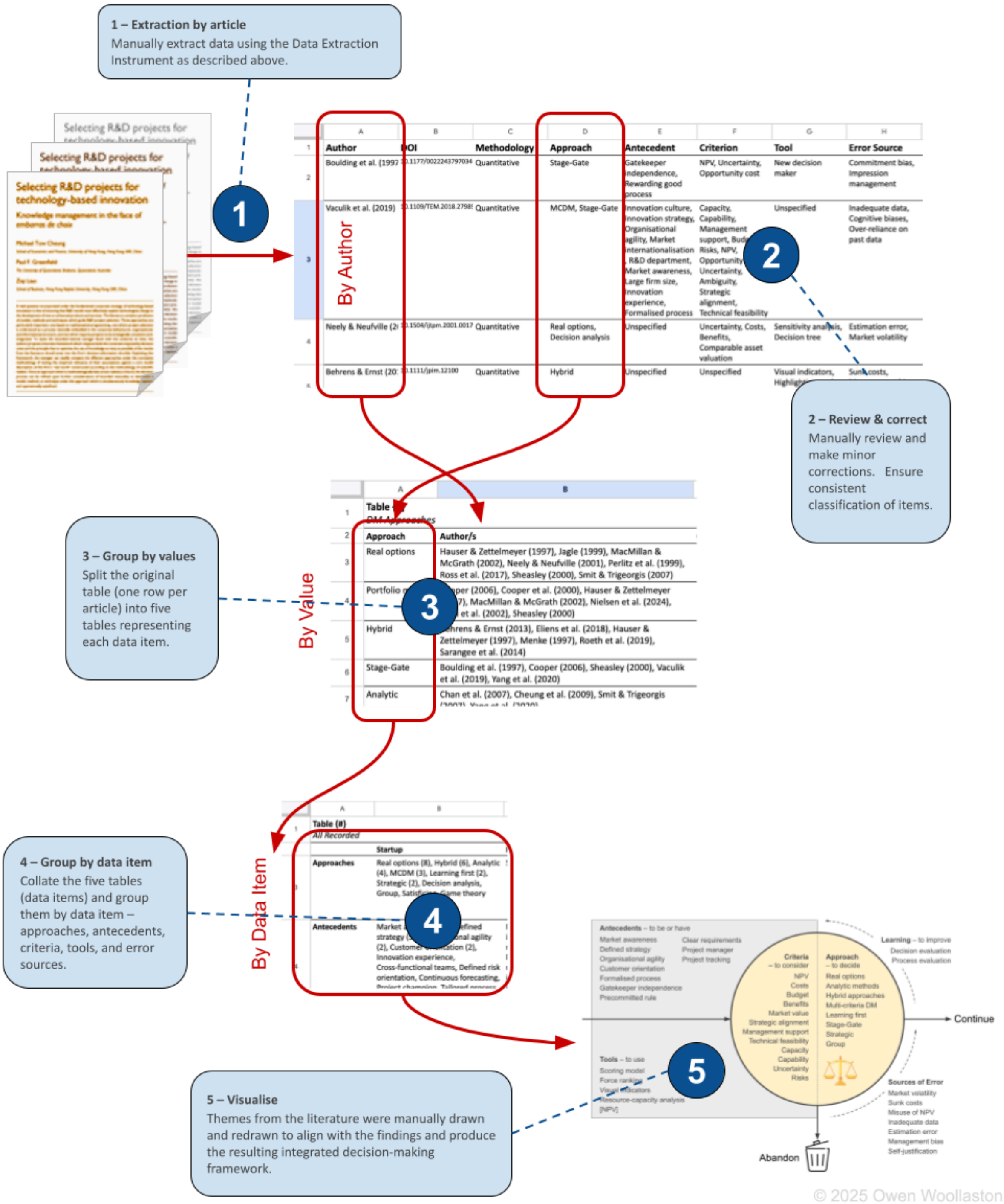
2.4.4 Data Transformation and Insights

With the data appropriately extracted and categorised, the approach described by Cronin and George (2023) was followed to integrate and present the results. The first attempt at integration began with visually tabulating the collected data items (decision-making factors) against their applicability to different organisational maturity levels. For example, some tools and approaches were suited to startups, while others were more effective in large enterprises. Next, the table was intuitively reorganised based on capabilities rather than maturity, providing a cleaner separation of data values. Organising by maturity or capabilities was subsequently abandoned because nearly all factors applied in one way or another to all organisations. Third, the layout was visually refined, tidied, and reorganised. The iterations of this process are in Appendix D. The output of this process formed the basis for the conceptual model presented in the discussion section that follows. Finally, as the results were reviewed and interpreted, key themes were identified in each of the categories and these were reworked into the framework. For example, having a defined strategy was an overarching theme in the antecedents category, so “defined strategy” is represented in the model. Thus, the framework was initially developed based on an intuitive organisation of the concepts and validated against a thematic review of literature findings.

Figure 9 captures an overview of the synthesis method. This method addresses the second aim of this thesis: To develop a decision-making framework. The following sections present and explore the results of the literature review and build toward development and presentation of the framework.

Figure 9

Summary of Synthesis Method



2.5 Literature Review Results

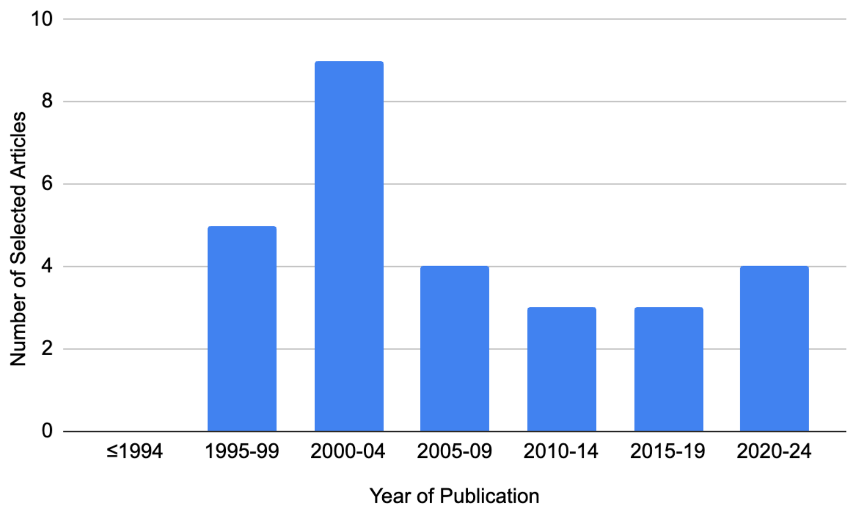
Following the method described in the previous section, the R&D decision-making was reviewed, and decision-making factors were extracted and synthesised. In the following sections, the results are presented, first with the overall state of the literature followed by results for each category – antecedents, approach, criteria, sources of error and tools.

2.5.1 Credibility of AI-Assisted Findings

Before presenting the findings, it is necessary to highlight one significant impact of using AI to rank articles by relevance. A common approach in traditional literature searches is to bias toward more recent literature. The underlying assumption is that recency is a good proxy for relevance. However, regarding this topic, the AI-assisted approach disagreed in many cases. Regardless of the publication date, the AI ranked articles against their conceptual relevance, leading to many turn-of-the-century articles being selected. Upon manual inspection, the AI findings were found to be more relevant, which is also empirically demonstrated in Appendix G. More recent works are generally found to be derivative and less relevant to the research question, and potentially relevant contemporary topics – like AI-assisted decision-making – have not yet been addressed in the context of high-tech NPD. The outcome is that the results presented here are oriented around foundational works, and the discussion highlights areas for future research.

Figure 10

Selected Articles by Year of Publication



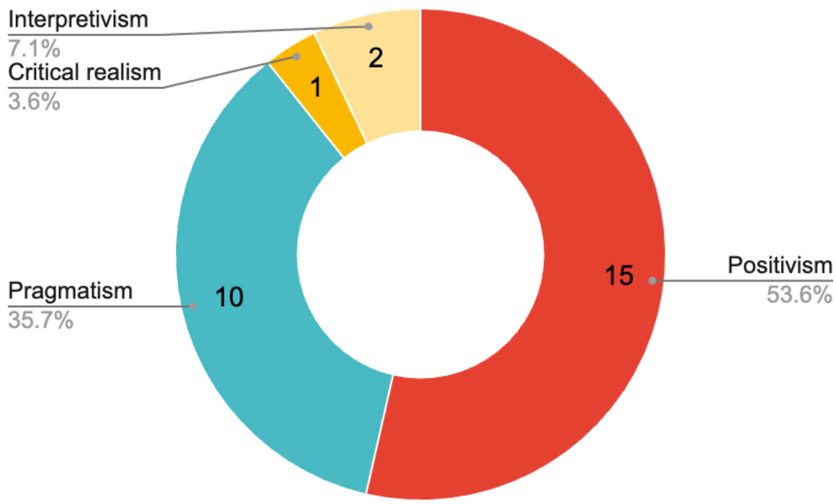
2.5.2 State of the Literature

Of 2877 articles screened, 28 met the inclusion criteria (Table 7, page 44) and were selected for data extraction and synthesis. All selected articles and extracted data is in Appendix C. Interest in go/no-go decisions in high-tech NPD peaked just after the turn of the century and plateaued after 2005, as indicated in Figure 10 above. The peak of interest correlates to the application of real options theory to R&D, some thirty years after it was established as a theory by Stewart Myers in 1977 (Myers, 1977).

The research philosophy of the selected articles was identified as shown in Figure 11. The predominant philosophy was positivism, which focuses on establishing the objective reality of the situation (usually quantitatively), as is the focus of real options theory, and perhaps reflecting the influence of rational economics on the field.

Figure 11

Research Philosophy of Included Articles



In terms of the methods used by these publications, thirteen were quantitative papers, ten qualitative and five applied mixed methods, indicating a balanced set of perspectives in this review – see Table 12.

Table 12

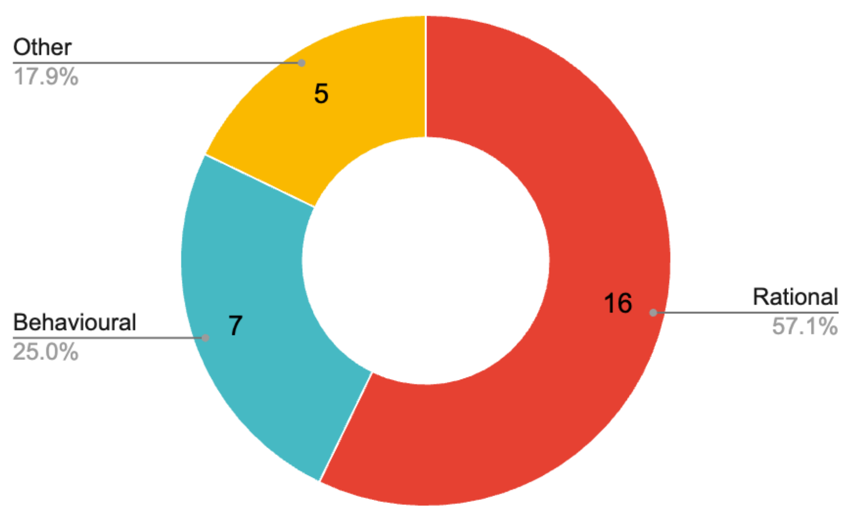
Selected Article Methodology

Methodology	Author/s	Count
Quantitative	Behrens & Ernst (2013), Boulding et al. (1997), Chan et al. (2007), Eliens et al. (2018), MacMillan & McGrath (2002), Menke (1997), Neely III & De Neufville (2001), Nielsen et al. (2024), Roeth et al. (2019), Ross et al. (2017), Schmidt et al. (2001), Vaculik et al. (2019), Yang et al. (2020)	13
Qualitative	Cheung et al. (2009), Cooper (2006), Cooper et al. (2000), Mikkola (2001), Perlitz et al. (1999), Pillai et al. (2002), Sarangee et al. (2014), Sheasley (2000), Wang et al. (2010), Zammar et al. (2023)	10
Mixed	Hauser & Zettelmeyer (1997), Jagle (1999), Linton et al. (2002), Loch (2000), Smit & Trigeorgis (2007)	5
TOTAL (3 distinct methodologies)		28

The articles were also classified by the branch of economics they used, as shown in Figure 12. In this figure, "Other" means the article wasn't based on economic theory.

Figure 12

Economic Lens of Included Articles



These results indicate the influence of positivist rational economics in the included articles, particularly evident in the data between 1995 and 2010. However, interest in behavioural economics in NPD has increased more recently with 71% (5 of 7) of the articles published in the decade since

2014. This overall dearth of behavioural-focused research, combined with the recent uplift of interest, might indicate more research is useful in this space.

While these philosophical dispositions frame the broader research landscape, a more detailed examination of decision-making factors shows the factors that shape R&D investment decisions. The following section categorises these into key themes to provide a structured review of the literature.

2.5.3 Overview of Decision-Making Factors

The review uncovered a bewildering array of decision-making approaches, antecedents, criteria, tools, and error sources – 126 unique factors, as shown in Figure 13 and summarised in Table 13 below.   Even when the values were classified to reduce complexity, 34 remained.

Figure 13

A Bewildering Array of Decision-Making Factors

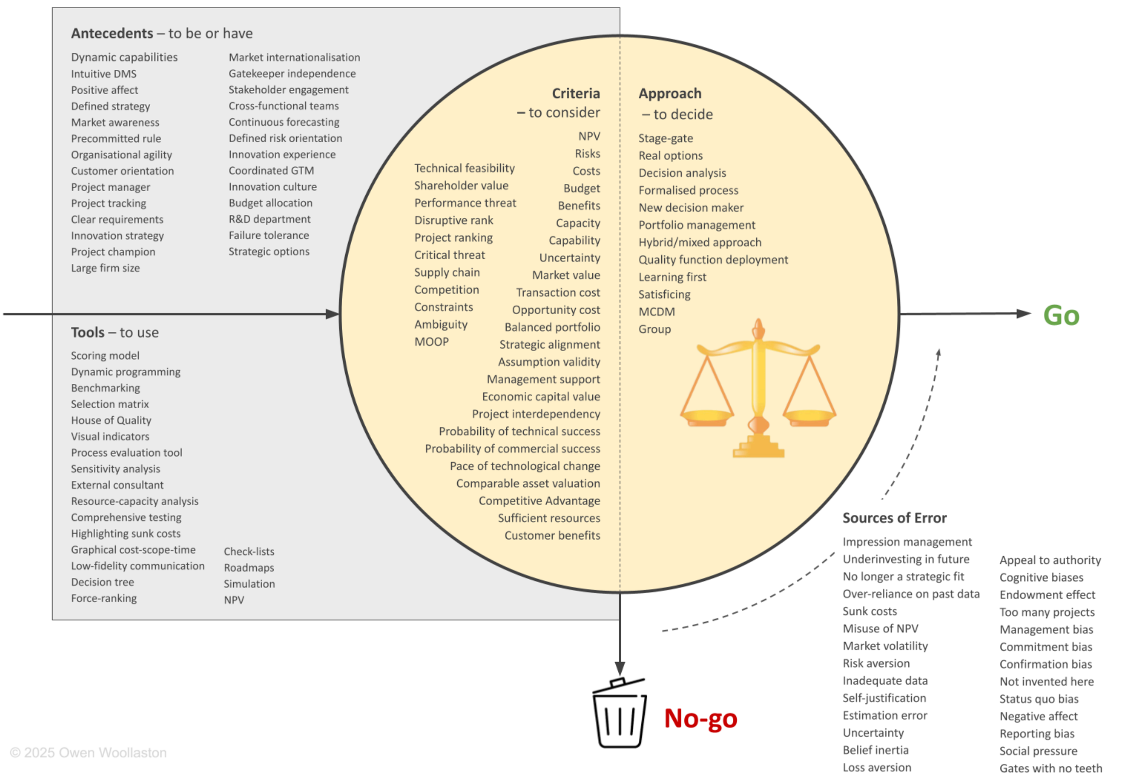


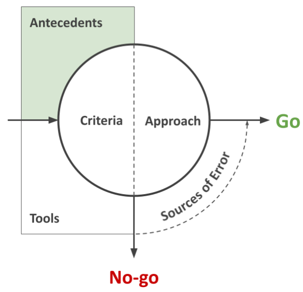
Table 13

Decision-Making Factor Tally

Data Item	Number of Classes	Number of Factors
Approaches	6	14
Antecedents	6	27
Criteria	12	35
Tools	6	23
Error Sources	4	27
TOTAL	**34**	**126**

As mentioned in the synthesis method, the 126 decision-making factors presented a challenge for integration, so these were further grouped into classes of manageable size. The decision-making factors were organised into five categories from the conceptual framework to resolve this complexity. To aid the reader in following the progress of the results in the context of the conceptual framework, each category is accompanied by a locator icon. This begins with the results for antecedents.

2.5.4 Decision-Making Antecedents



Antecedents are preconditions for decision-making, such as strategy and capabilities that must be established before investment decisions. They are something to be or have, occurring before the actual decision point.

2.5.4.1 Review Results and Classification: Antecedents

The review identified twenty-seven distinct antecedents (Table 14 and Figure 14 below), which were initially classified into six groups: culture, discipline, perspective, process, strategy, and structure. The most emphasised antecedents relate to strategic considerations, the most poignant and fundamental of which are having a defined strategy (Cooper, 2006; Cooper et al., 2000; Loch, 2000; MacMillan & McGrath, 2002; Menke, 1997) and an awareness of the market (Loch, 2000; Menke, 1997; Nielsen et al., 2024; Sarangee et al., 2014; Smit & Trigeorgis, 2007; Vaculik et al., 2019). A cluster of research orients around the use of stage-gates, particularly having pre-committed rules that commit managers to a course of action (go or no-go) contingent on specific criteria ahead of time (Cooper et al., 2000; Sarangee et al., 2014; Sheasley, 2000; Yang et al., 2020) and the use of independent gatekeepers who are not invested in the outcome and have an outsider's perspective (Boulding et al., 1997; Eliens et al., 2018; Roeth et al., 2019; Sarangee et al., 2014).

Figure 14

Decision-Making Antecedent Classes

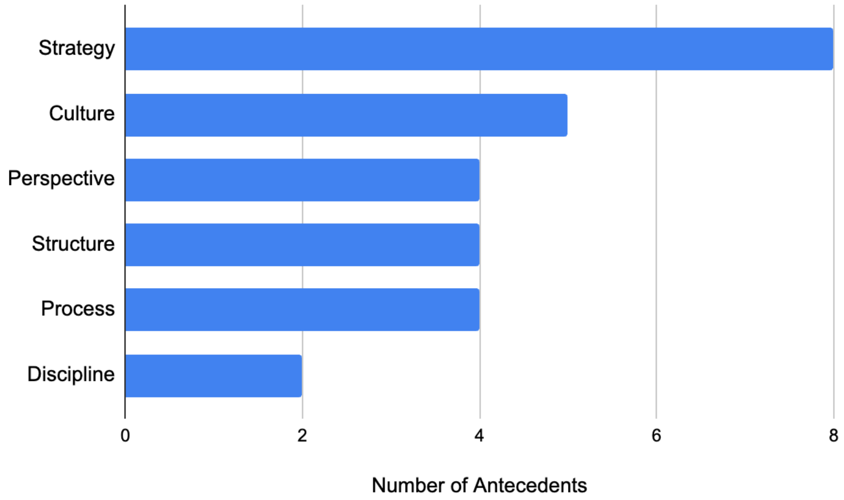


Table 14

Reported Decision-Making Antecedents

Class	Antecedent	Author/s
Strategy	Budget allocation, Defined risk orientation, Defined strategy, Dynamic capabilities, Innovation strategy, Market awareness, Market internationalisation, Strategic options	Chan et al. (2007), Cooper (2006), Cooper et al. (2000), Loch (2000), MacMillan & McGrath (2002), Menke (1997), Nielsen et al. (2024), Sarangee et al. (2014), Smit & Trigeorgis (2007), Vaculik et al. (2019)
Culture	Customer orientation, Innovation culture, Intuitive DMS, Organisational agility, Positive affect	Chan et al. (2007), Loch (2000), Menke (1997), Roeth et al. (2019), Vaculik et al. (2019)
Perspective	Failure tolerance, Gatekeeper independence, Innovation experience, Project champion	Boulding et al. (1997), Eliens et al. (2018), Loch (2000), Roeth et al. (2019), Sarangee et al. (2014), Vaculik et al. (2019)
Structure	Cross-functional teams, Large firm size, Project manager, R&D department	Loch (2000), Menke (1997), Vaculik et al. (2019)
Process	Continuous forecasting, Coordinated GTM, Project tracking, Stakeholder engagement	Menke (1997), Pillai et al. (2002), Sarangee et al. (2014)
Discipline	Clear requirements, Precommitted rule	Cooper et al. (2000), Menke (1997), Sarangee et al. (2014), Sheasley (2000), Yang et al. (2020)

2.5.4.2 Interpretation and Synthesis: Antecedents

The results, classification and literature highlighted several key themes. The first was the need for a **defined strategy**, which included factors such as product strategy, strategic budget allocation (Hutchison-Krupat & Kavadias, 2015), and identifying strategic options (Smit & Trigeorgis, 2007; Vaculik et al., 2019). In the context of go/no-go decision-making, the test of a good strategy is its ability to be tested against the go and no-go options to determine strategic alignment.

The second key theme was the importance of having the right **capabilities** in place. The literature identifies four (strategic) capabilities that should be developed in an organisation to support new product development. Organisational agility (Chan et al., 2007; Vaculik et al., 2019) and dynamic capabilities (Teece et al., 1997; Teirlinck & Spithoven, 2012) provide the organisation with the ability to embrace change and to pivot when change dictates or when necessitated by the evolving awareness of markets. Customer orientation (Loch, 2000; Menke, 1997) gives managers an empathetic perspective of the customer's wants, needs, roadmap and likely challenges, which can be translated into product and service offerings. More broadly, market awareness (Loch, 2000; Menke, 1997; Nielsen et al., 2024; Sarangee et al., 2014; Smit & Trigeorgis, 2007; Vaculik et al., 2019) is of paramount importance and often neglected or under-resourced in the organisations due to high (perceived or real) research costs. Finally and seemingly mundane is project tracking (Menke, 1997; Sarangee et al., 2014), without which it is difficult to gauge progress toward gates or against milestones in the contracts discussed above.

Contracting, third, is emphasised throughout the rhetoric of the literature and appears as a process of making explicit what is implicit and bringing to awareness what is assumed. Cooper et al. (2000), Menke (1997), Sarangee et al. (2014), Sheasley (2000) and Yang et al. (2020) all advocate establishing clear requirements and "precommiting" to decision-making rules and criteria well before resolve is tested and commitment escalation becomes possible.

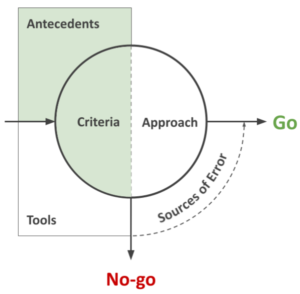
Good contracting is pivotal to R&D decision-making. In R&D, the stage-gate and other governance processes facilitate contracting, particularly the pre-agreement of go/no-go criteria. If, at the start of the development, a management team and stakeholders acknowledge the possibility of failure and agree on the no-go criteria, then stakeholder expectations and reactions are much easier to manage when executing the no-go option. While perhaps disappointing, it is more likely that managers' reputations remain intact. Thus, contracting in this context aims to manage expectations and identify and retain options that enable agility as uncertainty is progressively resolved through the stage-gate process. Up-front commitment to gate criteria and milestones will go a long way to alleviate impression management and reputation risk, thus allowing managers and stakeholders to make more objective choices along the way. As Sarangee et al. (2014) summarise:

"From a project-level perspective, decision-makers can de-escalate commitment from troubled product innovation projects by (1) defining measurable success and failure criteria at the beginning of the project; (2) defining the requirements of the project very clearly; (3) ensuring general firm-wide agreement upon these criteria and project requirements; (4) evaluating, reviewing, and tracking their projects regularly; and (5) cancelling the projects that do not meet the previously agreed-upon requirements."

The fourth theme is **oversight**. A project manager or champion supports the project by coordinating activities and reporting on progress, but that might not be enough to avoid commitment escalation. Eliens et al. (2018), Roeth et al. (2019) and Sarangee et al. (2014) take this one step further by advocating "gatekeeper independence" where the ultimate decision-making authority is not invested in the project. A product council or other governance groups might implement this.

With the right culture, strategy, oversight, and pre-agreed-upon go/no-go conditions, a firm is well positioned to make better go/no-go decisions. While these antecedents establish the necessary conditions for decision-making, the next crucial aspect involves the criteria used to evaluate potential investments. These criteria define the parameters against which decisions are assessed.

2.5.5 Decision-Making Criteria



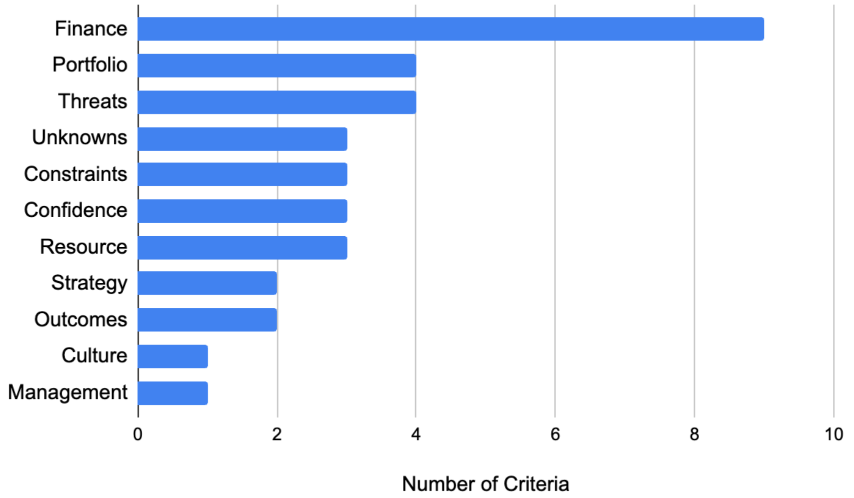
Criteria are the qualitative or quantitative factors considered when evaluating a decision, such as cost, risk, return, and strategic alignment. They are things to consider and can take many forms. They can be objective or subjective, quantitative or qualitative, categorical or binary. Still, they all represent an understanding of one aspect of the state of the R&D project, product, or environment.

2.5.5.1 Review Results and Classification: Criteria

Thirty-five decision-making criteria were identified through the review. These were grouped into confidence, constraints, error sources, external pressure, finance, outcomes, portfolio, resource, strategy, and threats, as shown in Figure 15 below and Table 15 (page 62). Unsurprisingly, financial criteria dominate as a class with net present value (Boulding et al., 1997; Jagle, 1999; MacMillan & McGrath, 2002; Nielsen et al., 2024; Perlitz et al., 1999; Pillai et al., 2002; Vaculik et al., 2019) and benefits (Cooper, 2006; Hauser & Zettelmeyer, 1997; Jagle, 1999; MacMillan & McGrath, 2002; Neely III & De Neufville, 2001; Nielsen et al., 2024; Perlitz et al., 1999) being the most cited criteria. Still, subjective assessments stand out when considering criteria such as ambiguity, risks, uncertainty, benefits, competitive advantage, strategic alignment, capability, probability of commercial success, probability of technical success, and technical feasibility. Many factors exist to consider, prioritise, and align with the strategy. Linton et al. (2002) also identified a long list of decision-making criteria (metrics in their terminology).

Figure 15

Decision-Making Criterion Classes



2.5.5.2 Interpretation and Synthesis: Criteria

Four criteria-related themes were apparent in the findings and classification above, as well as rhetoric in the literature. First, testing each option against its **strategic alignment** is critical to go/no-go decision making (Cooper, 2006; Cooper et al., 2000; Hauser & Zettelmeyer, 1997; Mikkola, 2001; Nielsen et al., 2024; Vaculik et al., 2019), in particular it is critical to maintain a view of whether or not the key benefits will be realised (Neely III & De Neufville, 2001; Jagle, 1999; Perlitz et al., 1999; Hauser & Zettelmeyer, 1997; Nielsen et al., 2024; Mikkola, 2001; MacMillan & McGrath, 2002; Cooper, 2006). If the realisation of strategic outcomes or project benefits is at risk, then the no-go option should be considered seriously. Further, if the opportunity cost (Boulding et al., 1997; Vaculik et al., 2019) incurred by marginalising another more attractive project would upset the portfolio balance (Cooper et al., 2000), then through no fault of the project delivery team, a portfolio manager should consider abandoning the project.

Table 15

Reported Decision-Making Criteria

Class	Criterion	Author/s
Finance	Budget, Comparable asset valuation, Costs, Economic capital value, MOOP, Market value, NPV, Shareholder value, Transaction cost	Boulding et al. (1997), Chan et al. (2007), Cooper et al. (2000), Hauser & Zettelmeyer (1997), Jagle (1999), MacMillan & McGrath (2002), Neely III & De Neufville (2001), Nielsen et al. (2024), Perlitz et al. (1999), Pillai et al. (2002), Sheasley (2000), Smit & Trigeorgis (2007), Vaculik et al. (2019)
Portfolio	Balanced portfolio, Disruptive rank, Opportunity cost, Project ranking	Boulding et al. (1997), Cooper et al. (2000), Hauser & Zettelmeyer (1997), MacMillan & McGrath (2002), Nielsen et al. (2024), Vaculik et al. (2019)
Threats	Critical threat, Competition, Pace of technological change, Performance threat	MacMillan & McGrath (2002), Yang et al. (2020)
Unknowns	Ambiguity, Risks, Uncertainty	Boulding et al. (1997), Jagle (1999), MacMillan & McGrath (2002), Neely III & De Neufville (2001), Perlitz et al. (1999), Pillai et al. (2002), Sheasley (2000), Vaculik et al. (2019)
Constraints	Constraints, Project interdependency, Supply chain	Chan et al. (2007), MacMillan & McGrath (2002), Nielsen et al. (2024)
Confidence	Probability of commercial success, Probability of technical success, Technical feasibility	Cooper (2006), MacMillan & McGrath (2002), Vaculik et al. (2019)
Resource	Capability, Capacity, Sufficient resources	Cooper et al. (2000), MacMillan & McGrath (2002), Menke (1997), Nielsen et al. (2024), Vaculik et al. (2019)
Strategy	Competitive Advantage, Strategic alignment	Cooper (2006), Cooper et al. (2000), Hauser & Zettelmeyer (1997), Mikkola (2001), Nielsen et al. (2024), Vaculik et al. (2019)
Outcomes	Benefits, Customer benefits	Cooper (2006), Hauser & Zettelmeyer (1997), Jagle (1999), MacMillan & McGrath (2002), Mikkola (2001), Neely III & De Neufville (2001), Nielsen et al. (2024), Perlitz et al. (1999)
Culture	Management support	Loch (2000), MacMillan & McGrath (2002), Menke (1997), Sheasley (2000), Vaculik et al. (2019)
Management	Assumption validity	Pillai et al. (2002), Sarangee et al. (2014)

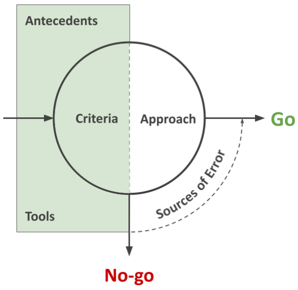
Second, numerous **financial** criteria were uncovered in the literature. They include the project budget and cost, and expect returns usually expressed as net present value (Boulding et al., 1997; Vaculik et al., 2019; Jagle, 1999; Perlitz et al., 1999; Nielsen et al., 2024; Pillai et al., 2002; MacMillan & McGrath, 2002). Maximum out-of-pocket (MOOP) is an important criterion for indicating the worst-case impact on cashflow and capital resourcing (Nielsen et al., 2024), and there are many other useful financial criteria as defined in Table 15 above.

Third, it is critical to understand the epistemology of the decision – what is known, and more importantly **what is unknown**. An understanding of risks (Vaculik et al., 2019; Jagle, 1999; Perlitz et al., 1999; MacMillan & McGrath, 2002) and uncertainty (Boulding et al., 1997; Vaculik et al., 2019; Neely III & De Neufville, 2001; Perlitz et al., 1999; Pillai et al., 2002; Sheasley, 2000) are also critical to NPD decision-making. Potential sources of ambiguity (Vaculik et al., 2019) and the validity of assumptions (Pillai et al., 2002; Sarangee et al., 2014) should also be understood. Technical feasibility (Vaculik et al., 2019; MacMillan & McGrath, 2002) is often unknown at the start of a project and must be progressively elaborated as the project proceeds. External threats (MacMillan & McGrath, 2002; Yang et al., 2020) also add a degree of risk to the project that must be understood. By their nature, and as indicated in the literature, R&D projects face many unknowns.

Finally, the firm should have a good understanding of **resources** at its disposal and the impact of the go/no-go decision on them. Does the firm have capacity and the capability (Vaculik et al., 2019; Nielsen et al., 2024; MacMillan & McGrath, 2002; Cooper et al., 2000) to execute the work effectively? And importantly, does the team have the full backing of senior management? (Loch, 2000; MacMillan & McGrath, 2002; Menke, 1997; Sheasley, 2000; Vaculik et al., 2019). Without the right resources, the project is likely to fail.

However, having the right criteria alone is insufficient; decision-makers also require tools to analyse and interpret these criteria. The next section explores the decision-making tools identified in the literature that facilitate structured evaluation and improve decision accuracy.

2.5.6 Decision-Making Tools



Tools are mental or physical aids that enhance decision-making by measuring, analysing, and interpreting critical information. They are to use and inform. Like antecedents, tools should be in place before the decision point. Often tools are used to process information, which needs to be collated and processed to provide insight. It can be difficult and time-consuming to do this in the moment.

2.5.6.1 Review Results and Classification: Tools

From the literature, twenty-three decision-making tools were identified in the R&D go/no-go decisions literature. These were classified into analytic, management, multi-criteria, process, psychological, and strategy, as shown in Figure 16 and Table 16 below. The single most commonly identified tool was a scoring model (belonging to the multi-criteria class), in which criteria are weighed against requirements to facilitate decision-making and clarification of priorities (Cooper, 2006; Cooper et al., 2000; Pillai et al., 2002; Sheasley, 2000; Wang et al., 2010).

Figure 16

Decision-Making Tool Classes

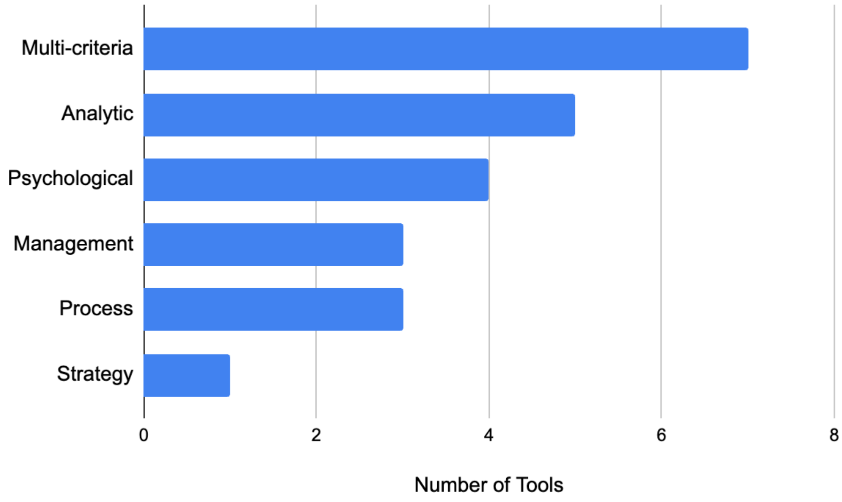


Table 16

Reported Decision-Making Tools

Class	Tool	Author/s
Multi-criteria	Balanced scorecard, Force-ranking, Game theory, House of Quality, Scoring model, Selection Matrix, Selection matrix	Cooper (2006), Cooper et al. (2000), Mikkola (2001), Pillai et al. (2002), Sheasley (2000), Smit & Trigeorgis (2007), Wang et al. (2010)
Analytic	Decision tree, Dynamic programming, NPV, Sensitivity analysis, Simulation	Chan et al. (2007), Jagle (1999), Neely III & De Neufville (2001), Nielsen et al. (2024), Smit & Trigeorgis (2007)
Psychological	Graphical cost-scope-time, Highlighting sunk costs, Low-fidelity communication, Visual indicators	Behrens & Ernst (2013), Linton et al. (2002), Pillai et al. (2002), Schmidt et al. (2001)
Management	External consultant, Process evaluation tool, Resource-capacity analysis	Behrens & Ernst (2013), Cheung et al. (2009), Cooper et al. (2000), MacMillan & McGrath (2002)
Process	Check-lists, Comprehensive testing, Roadmaps	Cooper et al. (2000), Sarangee et al. (2014)
Strategy	Benchmarking	Sarangee et al. (2014)

A diverse range of tools was identified. Analytical tools such as sensitivity analysis (Jagle, 1999; Neely III & De Neufville, 2001; Smit & Trigeorgis, 2007), simulation (Cooper, 2006; Cooper et al., 2000), and decision trees (Jagle, 1999; Neely III & De Neufville, 2001) were frequently referenced. Multi-criteria tools (mostly variants on prioritisation and weighting tools) were also commonly reported, as were psychological tools – mainly concerned with presenting information – and a miscellany of management tools. Interestingly, although strategic conditions dominated as an antecedent, few strategic tools were cited, namely benchmarking and roadmaps (Sarangee et al., 2014).

A range of tools have been developed to alleviate uncertainty and ambiguity. These can be loosely categorised as subjective or objective. Subjective tools include scoring models where options are assigned weighted values. These provide a way for multiple participants to contribute independently (Cooper, 2006; Cooper et al., 2000; Pillai et al., 2002; Sheasley, 2000; Wang et al., 2010). These can be presented as selection matrices in which multiple options can be compared (Mikkola, 2001). Criteria abound in R&D decision-making and are useful in identifying areas of uncertainty that need exploring. Another popular subjective tool is force ranking (Cooper, 2006; Cooper et al., 2000), in which participants must rank projects from most preferred to least preferred, and there can be no ties. This forces participants to consider the key differences between projects and negotiate priority conflicts.

Objective tools involve measurement and analysis. Included are decision trees (aka option trees) and benchmarking, where managers assess potential products against their own and competitors' products (Sarangee et al., 2014). Simulation (Chan et al., 2007; Smit & Trigeorgis, 2007) can be a powerful tool for evaluating future possibilities beyond the probabilistic mathematics of decision trees, but they can also instil false confidence and should be used alongside sensitivity analysis so that managers are aware of how the input parameters and assumptions affect the potential outcomes (Jagle, 1999; Neely III & De Neufville, 2001; Smit & Trigeorgis, 2007). Finally, resource-capacity analysis is essential to ascertain whether or not the project will overcommit the firm and result in diluted outcomes (Cooper et al., 2000; MacMillan & McGrath, 2002).

2.5.6.2 Interpretation and Synthesis: Tools

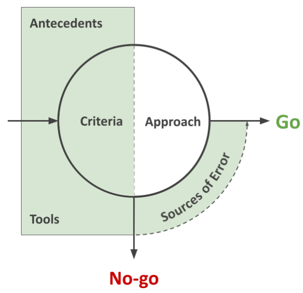
Three themes emerged regarding R&D tools. The first involved **planning and communicating** plans involving product roadmaps (Sarangee et al., 2014) and other visual indicators (Linton et al., 2002) of progress to illustrate the overarching plan for a product, product line, or product family. These tools help position the project within a broader objective and clarify decisions regarding timing, progress, and the long road to product success.

The second theme was around **options analysis**, in which tools are used to evaluate options available to the project team. These are usually analytical and include decision trees (Neely III & De Neufville, 2001; Jagle, 1999), simulation (Chan et al., 2007; Smit & Trigeorgis, 2007), benchmarking against competitors' products (Sarangee et al., 2014), sensitivity analysis (Neely III & De Neufville, 2001; Jagle, 1999; Smit & Trigeorgis, 2007) and resource-capacity analysis (MacMillan & McGrath, 2002; Cooper et al., 2000). These analytical tools help managers understand the impact of each potential course of action.

The final theme was around **portfolio review** tools. These support managers to answer a bigger question: Is this still the right product to develop? They evaluate products (or potential products) against one another, and include scoring models and selection matrices (Pillai et al., 2002; Sheasley, 2000; Cooper, 2006; Wang et al., 2010; Cooper et al., 2000) such as multi-criteria decision-making, and force-ranking (Cooper, 2006; Cooper et al., 2000) in which projects are ranked in a single list.

While decision-making tools enhance objectivity and structure, their effectiveness depends on how well they mitigate common errors. The following section examines sources of error that can distort decision outcomes, such as cognitive biases, data inaccuracies, and mismanagement.

2.5.7 Sources of Error in Decision-Making



Sources of Error are factors that knowingly or unknowingly introduce bias or inaccuracies into the decision-making process. They are things to mitigate or avoid. As discussed in the theoretical foundations, sources of error such as overconfidence bias (Eliens et al., 2018) can lead to commitment escalation where managers continue investing in a project the firm would be objectively better off abandoning.

2.5.7.1 Review Results and Classification: Sources of Error

Twenty-seven distinct sources of error in decision-making were identified in the literature. These were grouped into bias, change, mismanagement, and uncertainty, as shown in Figure 17 and Table 17 below. Bias accounts for two-thirds of all reported error sources in R&D decision-making, with little focus on any particular bias. Table 17 lists biases as well as other sources of error, most notably internal and external change (Neely III & De Neufville, 2001; Pillai et al., 2002; Smit & Trigeorgis, 2007), uncertainty in data and estimates (Cooper et al., 2000; Neely III & De Neufville, 2001; Nielsen et al., 2024; Ross et al., 2017; Smit & Trigeorgis, 2007; Vaculik et al., 2019) and types of mismanagement including misuse of NPV (Cooper, 2006; Cooper et al., 2000; Jagle, 1999) and running too many projects (Cooper et al., 2000; MacMillan & McGrath, 2002).

Figure 17

Decision-Making Error Classes

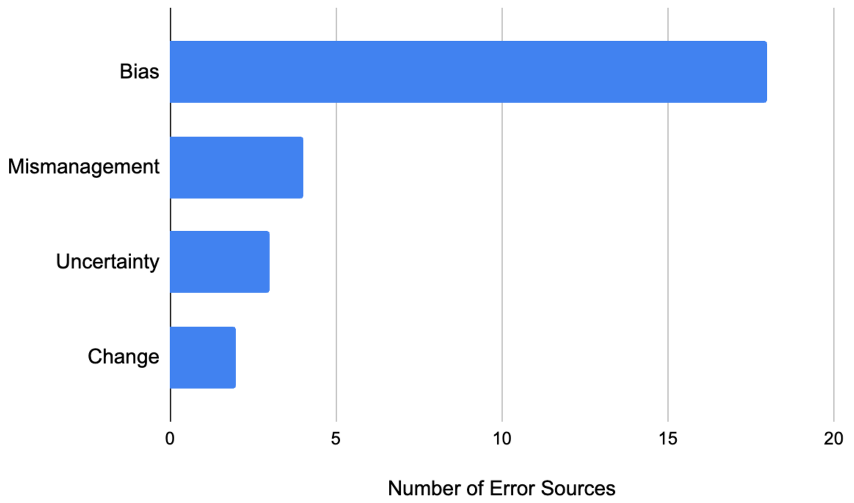


Table 17

Reported Decision-Making Error Sources

Class	Error Source	Author/s
Bias	Appeal to authority, Belief inertia, Cognitive biases, Commitment bias, Confirmation bias, Endowment effect, Impression management, Loss aversion, Management bias, Negative affect, Not invented here, Over-reliance on past data, Reporting bias, Risk aversion, Self-justification, Social pressure, Status quo bias, Sunk cost fallacy	Behrens & Ernst (2013), Boulding et al. (1997), Cooper (2006), Eliens et al. (2018), Hauser & Zettelmeyer (1997), Nielsen et al. (2024), Roeth et al. (2019), Sarangee et al. (2014), Schmidt et al. (2001), Vaculik et al. (2019), Yang et al. (2020)
Mismanagement	Gates with no teeth, Misuse of NPV, Too many projects, Underinvesting in future	Cooper (2006), Cooper et al. (2000), Jagle (1999), MacMillan & McGrath (2002), Yang et al. (2020)
Uncertainty	Estimation error, Inadequate data, Uncertainty	Cooper et al. (2000), Neely III & De Neufville (2001), Nielsen et al. (2024), Ross et al. (2017), Smit & Trigeorgis (2007), Vaculik et al. (2019)
Change	Market volatility, No longer a strategic fit	Neely III & De Neufville (2001), Pillai et al. (2002), Smit & Trigeorgis (2007)

2.5.7.2 Interpretation and Synthesis: Sources of Error

Four major themes emerged regarding sources of error in R&D decision-making, aligning directly to the classification of the sources of error in Table 17 above. The first was (psychological) **bias**, which occurs in many forms. A formidable list that all contribute to commitment escalation are sunk cost fallacy (Behrens & Ernst, 2013; Roeth et al., 2019; Yang et al., 2020), self-justification (Eliens et al., 2018; Yang et al., 2020), commitment bias (Boulding et al., 1997), and impression management (Boulding et al., 1997); belief inertia, loss aversion, the endowment effect, status quo bias, and confirmation bias, all cited by Eliens et al. (2018); and finally reporting bias (Sarangee et al., 2014), "not invented here" (Hauser & Zettelmeyer, 1997) and social pressure (Schmidt et al., 2001). Biases that tend to negate commitment escalation are few, and include risk aversion (Cooper, 2006; Hauser & Zettelmeyer, 1997) and negative affect (Roeth et al., 2019).

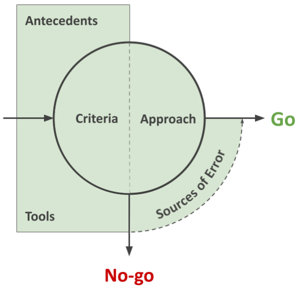
The second theme originates from **data quality**, which can introduce significant errors in the decision making process. Commonly, cost estimates can be wrong (Neely III & De Neufville, 2001; Smit & Trigeorgis, 2007) and market predictions suffer from inadequate data (Cooper et al., 2000; Vaculik et al., 2019) and high degrees of uncertainty (Nielsen et al., 2024; Ross et al., 2017). Also, numerical biases can sometimes arise from an overreliance on past data (Vaculik et al., 2019).

The third theme is **change**. Markets can be volatile, and change during the development of a product (Neely III & De Neufville, 2001; Pillai et al., 2002; Smit & Trigeorgis, 2007), or the firm might decide to change strategy making the project no longer a strategic fit (Pillai et al., 2002).

Fourth and finally, **mismanagement** can lead to large errors. Researchers find that an overreliance on net present value (Cooper, 2006; Cooper et al., 2000; Jagle, 1999) can lead to underrecognition of the importance of other factors. Attempting too many projects (and thus diluting resources) is a common issue (Cooper et al., 2000; MacMillan & McGrath, 2002). And finally, "gates with no teeth" (Yang et al., 2020) colourfully hints that go/no-go gates must be meaningfully applied to be effective.

These errors, caused by bias, data quality, change, and mismanagement, perturb the decision-making process, making it more difficult for managers to identify the best option. Recognising and mitigating them is essential, but decision-makers must also adopt a structured approach to ensure consistency and rigor in their evaluations. The next section examines different decision-making approaches, ranging from formal analytical methods to intuitive and social strategies.

2.5.8 Decision-Making Approaches



The final set of results entails decision-making approaches, which are the methodology used in decision-making. Shaped by philosophy, experience, and preferences, they can be formal or informal, analytic or intuitive. They are the *how* in the decision-making process.

2.5.8.1 Review Results and Classification: Approach

Fourteen distinct decision-making approaches were identified through the review. These were grouped into analytic, multi-criteria, portfolio, process, social, and strategic, as shown in Figure 18 and Table 18 below. The highest emphasis went on hybrid methods (Behrens & Ernst, 2013; Eliens et al., 2018; Hauser & Zettelmeyer, 1997; Loch, 2000; Menke, 1997; Roeth et al., 2019; Sarangee et al., 2014). Other leading approaches are the use of a stage-gate process and holding gate review meetings (Boulding et al., 1997; Cooper, 2006; Sheasley, 2000; Vaculik et al., 2019; Yang et al., 2020), portfolio management (Cooper, 2006; Cooper et al., 2000; Hauser & Zettelmeyer, 1997; MacMillan & McGrath, 2002; Nielsen et al., 2024; Pillai et al., 2002; Sheasley, 2000) or other formalised processes (Cheung et al., 2009; Cooper et al., 2000; Eliens et al., 2018; Loch, 2000; Mikkola, 2001; Sarangee et al., 2014; Vaculik et al., 2019). Real options featured highly as an analytic approach (Hauser & Zettelmeyer, 1997; Jagle, 1999; MacMillan & McGrath, 2002; Neely III & De Neufville, 2001; Perlitz et al., 1999; Ross et al., 2017; Sheasley, 2000; Smit & Trigeorgis, 2007). Overwhelmingly, a theme of formality of process emerges from the literature, with minimal recognition of social or strategic approaches.

Figure 18

Decision-Making Approach Classes

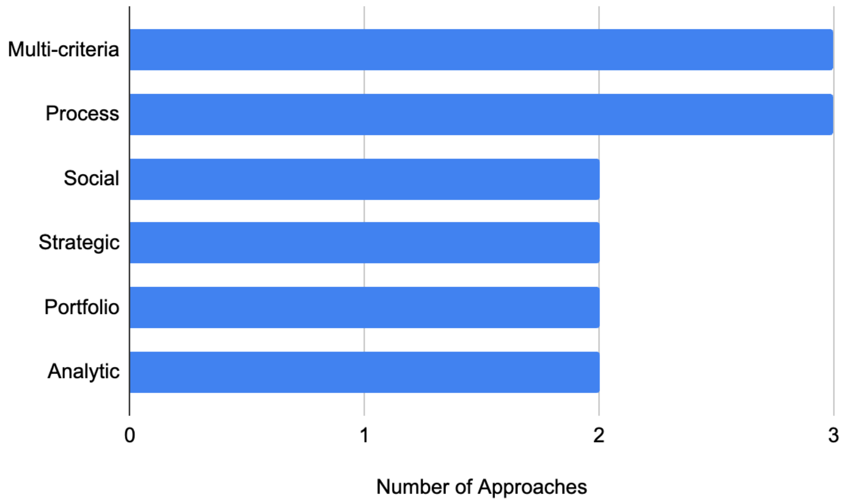


Table 18

Reported Decision-Making Approaches

Class	Approach	Author/s
Multi-criteria	Hybrid/mixed approach, MCDM, Quality function deployment	Behrens & Ernst (2013), Eliens et al. (2018), Hauser & Zettelmeyer (1997), Linton et al. (2002), Loch (2000), Menke (1997), Roeth et al. (2019), Sarangee et al. (2014), Vaculik et al. (2019), Wang et al. (2010), Zammar et al. (2023)
Process	Formalised process, Learning first, stage-gate	Boulding et al. (1997), Cheung et al. (2009), Cooper (2006), Cooper et al. (2000), Eliens et al. (2018), Loch (2000), Mikkola (2001), Ross et al. (2017), Sarangee et al. (2014), Sheasley (2000), Vaculik et al. (2019), Yang et al. (2020)
Social	Group, New decision maker	Boulding et al. (1997), Schmidt et al. (2001)
Strategic	Decision effectiveness, Satisficing	Cheung et al. (2009), Pillai et al. (2002)
Portfolio	Data envelopment analysis, Portfolio management	Cooper (2006), Cooper et al. (2000), Hauser & Zettelmeyer (1997), Linton et al. (2002), MacMillan & McGrath (2002), Nielsen et al. (2024), Pillai et al. (2002), Sheasley (2000)
Analytic	Decision analysis, Real options	Hauser & Zettelmeyer (1997), Jagle (1999), MacMillan & McGrath (2002), Neely III & De Neufville (2001), Perlitz et al. (1999), Ross et al. (2017), Sheasley (2000), Smit & Trigeorgis (2007)

2.5.8.2 Interpretation and Synthesis: Approach

At the forefront of the literature was an emphasis on the need for a **formal process** (Vaculik et al., 2019; Mikkola, 2001; Cheung et al., 2009; Loch, 2000; Sarangee et al., 2014; Cooper et al., 2000; Eliens et al., 2018). Formality is not clearly defined in the literature, but the essential concept expressed was the need for a pre-agreed process to guide the decision-making. Pre-agreement on the process is important because it establishes clear terms of engagement between managers when decision-making becomes difficult (Eliens et al., 2018).

The antecedents and criteria sections have already established the need to make strategically aligned decisions. However, managers' ability to evaluate options objectively is influenced by their level of commitment escalation and other biases. To achieve this, researchers have recommended many strategies based on one of two concepts: viewing the decision through **multiple lenses** (perspectives) and applying **distancing strategies** to achieve better objectivity.

There are many available lenses, including real options (Hauser & Zettelmeyer, 1997; Jagle, 1999; MacMillan & McGrath, 2002; Neely III & De Neufville, 2001; Perlitz et al., 1999; Ross et al., 2017; Sheasley, 2000; Smit & Trigeorgis, 2007), portfolio management (Cooper, 2006; Cooper et al., 2000;

Hauser & Zettelmeyer, 1997; MacMillan & McGrath, 2002; Nielsen et al., 2024; Pillai et al., 2002; Sheasley, 2000), multi-criteria decision-making (MCDM) (Linton et al., 2002; Vaculik et al., 2019; Zammar et al., 2023), and quality function deployment (Wang et al., 2010). These disciplines and theories provide different philosophies and methods to inform decisions. For example, real options theory takes a Bayesian (statistical) approach to evaluating options. In contrast, MCDM takes the approach of tabulating and considering many different qualitative and quantitative criteria, whereas quality function deployment takes a customer-centered cross-functional approach. No one lens provides the whole picture, but combined, they support the decision-makers in considering multiple perspectives, which is one effective way of distancing from the decision.

Distancing strategies are ways of removing oneself emotionally from a decision, and thus allowing for more objectivity. They are a potential antidote to commitment escalation, come in many forms, and may sometimes overlap with perspective taking (lenses). At one extreme the decision itself is passed over to an independent gatekeeper (Boulding et al., 1997; Eliens et al., 2018; Roeth et al., 2019; Sarangee et al., 2014) or group such as a product council or governance group. In another (social) form group decision-making (Boulding et al., 1997; Schmidt et al., 2001) encourages managers to share perspectives and challenge each other. In a third (temporal) form managers pre-commit to decision-making rules (Cooper et al., 2000; Sarangee et al., 2014; Sheasley, 2000; Yang et al., 2020) so that they are obligated to follow through if emotionally challenged later on.

By following the themes suggested in the literature – using a formal process, multiple lenses and distancing strategies – managers should find themselves making more objective and more beneficial decisions.

Taken together, the five categories of factors provide a comprehensive view of decision-making in high-tech R&D from the extant literature. The next section synthesises these insights into a broader framework, discussing their implications for research and practice while identifying opportunities for further study.

2.6 Discussion

This review of the state of the literature regarding R&D go/no-go decisions with the assistance of AI provides a nuanced collation on this topic. However, as shown in Figure 13 (page 56), they offer little practical guidance for SMEs, as the 126 identified factors would overwhelm decision-makers. A drawback of focused studies is that their results are published in discrete quanta, leaving time-poor practitioners to handle much of the interpretation and integration. This high integration effort exacerbates the research-practice gap (Bansal et al., 2012). A key aim of this thesis is to develop an integrated framework with practical value for managers. To achieve this, the model must be distilled to its essential components and structured holistically. To address this challenge, the next section summarises the results and introduces a structured go/no-go framework, synthesising key factors into a practical tool for R&D managers.

2.6.1 Go/No-Go Framework

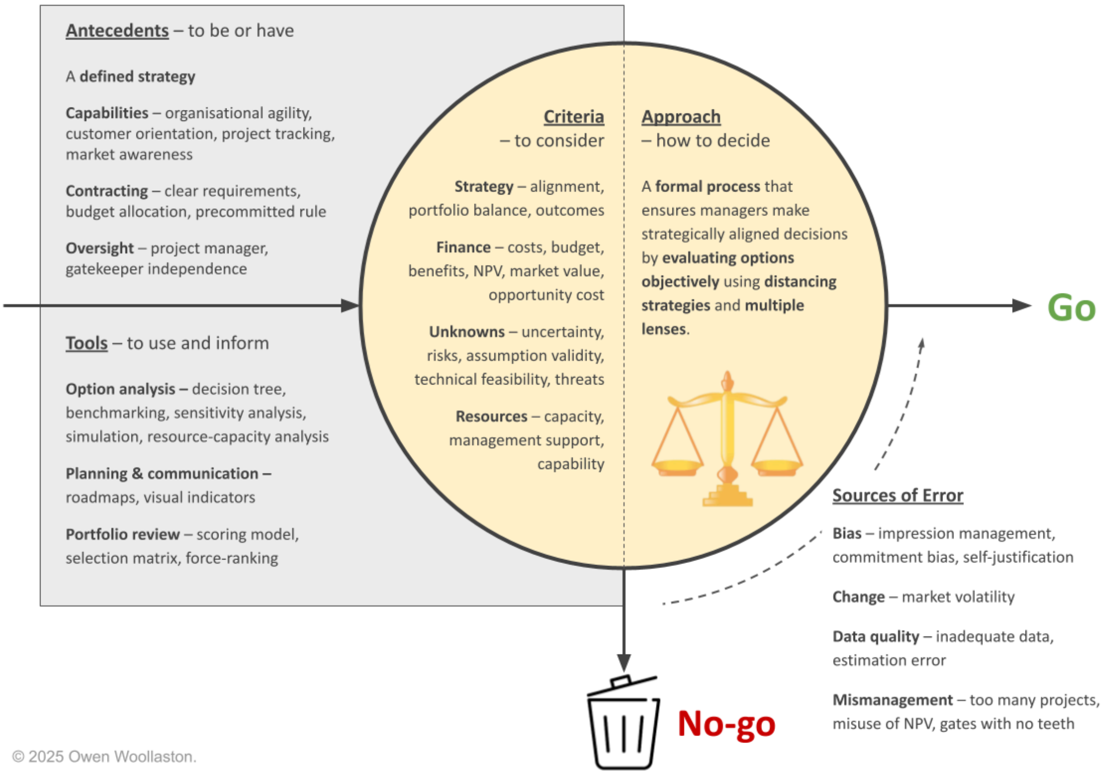
In presenting the results, the categories derived from the theoretical constructs were used to frame the decision-making factors, and essential factors (key groups of factors) were introduced, and are summarised here:

- ***Antecedents*** included a *defined strategy*, *capabilities* (organisational agility, customer orientation, project tracking, market awareness), *contracting* (clear requirements, budget allocation, precommitted rule), and *oversight* (project manager, gatekeeper independence).
- ***Criteria*** to consider included *strategy* (alignment, portfolio balance, outcomes), *finance* (costs, budget, benefits, NPV, market value, opportunity cost), *unknowns* (uncertainty, risks, assumption validity, technical feasibility, threats), and *resources* (capacity, management support, capability).
- ***Tools*** to use and inform included *planning and communication* (roadmaps, visual indicators), *option analysis* (decision tree, benchmarking, sensitivity analysis, simulation, resource-capacity analysis), and *portfolio review* (scoring model, selection matrix, force-ranking).
- ***Sources of error*** included *bias* (impression management, commitment bias, self-justification, sunk cost fallacy), *change* (market volatility), *data quality* (inadequate data, estimation error), and *mismanagement* (too many projects, misuse of NPV, weak gatekeeping).
- Finally, the distilled ***approach*** from the literature was summarised as: a *formal* process ensuring managers make strategically aligned decisions by evaluating options objectively using *distancing strategies and multiple lenses*.

By following the synthesis method, and applying the skeletal framework developed in the theoretical foundations (Figure 4, page 72) to the categories and essential factors, the proposed go/no-go framework ("Go/No-Go Framework" or "Framework") was produced and is depicted in Figure 19.

Figure 19

The Go/No-Go Framework



Unsurprisingly, strategy emerged as an overarching element of the model. In this context, strategy primarily refers to the product development strategy (not the broader organisational strategy) as this is more relevant to go/no-go decision-making. A product strategy primarily includes a development roadmap (representing strategic intent) against which go/no-go decisions can be tested. A key feature of any strategy is its testability, ensuring decision alignment and facilitating decision-making at all levels.

The Framework's approach, presented in sentence form, conveys the idea that there is one proposed approach with multiple features – formality, strategic alignment, multiple lenses and distancing. Thus, a direct answer to the research question "How should firms evaluate whether to continue or abandon projects in high-tech R&D?" was arrived at. Firms should follow: "A formal process that ensures managers make strategically aligned decisions by evaluating options objectively using distancing strategies and multiple lenses." This encapsulates the essential features of a good decision-making approach highlighted in the literature (Boulding et al., 1997; Cheung et al., 2009; Cooper, 2006; Cooper et al., 2000; Eliens et al., 2018; Loch, 2000; Mikkola, 2001; Ross et al., 2017; Sarangee et al., 2014; Sheasley, 2000; Vaculik et al., 2019; Yang et al., 2020).

Beyond providing structured decision guidance, the Framework can also be viewed through a risk-management lens as it seeks to mitigate key risks in R&D decision-making. The next section explores this perspective.

2.6.2 Conceptual Test of The Framework: A Risk Management Perspective

One function of the Framework is to mitigate the risk of suboptimal decisions by reducing bias and other factors that lead to commitment escalation. From this perspective, it can be viewed as a risk management tool focused on avoiding or minimising commitment escalation by guiding managers to make sound go/no-go decisions. As a risk management tool, the Framework can be tested conceptually by restructuring it in terms of risks and mitigations. Table 19 presents this by treating sources of error and unknowns as risks and relating them to targeted risk controls (tools and antecedents) and broader mitigations (strategy and approach). This makes explicit the Framework's implicit risk management structure.

Table 19

Go/No-Go Framework as a Risk Management Tool

<table>
<tr><th>Risk
(Error Sources and Criteria)</th><th>Risk Controls – Targeted
(Tools and Antecedents)</th><th>Risk Controls – Broad
(Strategy & Approach)</th></tr>
<tr><td>Mismanagement – too many projects, misuse of NPV, gates with no teeth</td><td rowspan="2">Contracting – clear requirements, budget allocation, precommitted rule
Oversight – project manager, gatekeeper independence
Portfolio review – scoring model, selection matrix, force-ranking</td><td>A defined strategy</td></tr>
<tr><td>Bias – impression management, commitment bias, self-justification</td><td>Planning & communication – roadmaps, visual indicators</td></tr>
<tr><td>Data quality – inadequate data, estimation error</td><td rowspan="3">Capabilities – organisational agility, customer orientation, project tracking, market awareness
Option analysis – decision tree, benchmarking, sensitivity analysis, simulation, resource-capacity analysis</td><td rowspan="3">Approach – A formal process that ensures managers make strategically aligned decisions by evaluating options objectively using distancing strategies and multiple lenses.</td></tr>
<tr><td>Change – market volatility</td></tr>
<tr><td>Unknowns – uncertainty, risks, assumption validity, technical feasibility, threats</td></tr>
</table>

From this perspective, risks from data issues, change, and unknowns are mitigated by key organisational capabilities and options thinking. Biases and mismanagement are mitigated by governance oversight, portfolio review and strong contracting (in the sense of pre-agreed decision process and criteria). A second tier of risk control comes from a clearly defined strategy, clear communication and a formal decision-making process. The risk management perspective helps us understand why particular features of the Framework are important. The plethora of biases (which each culminate in under- or overcommitment) is managed through project oversight in the form of transparent project management and reporting. Further, gatekeeper independence, which may take the form of internal or external governance (Garg, 2020), implements distancing and evaluation of options through the use of scoring models and selection matrices, which implement perspective taking. Yang et al. (2020) conclude, “A key component of a well-managed NPD project is [the skilful] management of stage-gate reviews”, and Sarangee et al. (2014) add, "One especially promising avenue seems to be the separation of decision-makers over repeated continuation decisions" – further emphasising the importance of distancing strategies.

While the Framework functions as a risk management tool, its reliance on a limited pool of behavioural economics research highlights critical gaps in the literature.

Table 20

Complete List of Included Behavioural Articles

Author	Title
Boulding et al. (1997)	Pulling the Plug to Stop the New Product Drain (Analyses the Extent of Senior Management's Commitment to Failing Product Launches)
Schmidt et al. (2001)	New Product Development Decision-Making Effectiveness: Comparing Individuals, Face-to-Face Teams, and Virtual Teams
Behrens & Ernst (2013)	What Keeps Managers Away From a Losing Course of Action? Go/Stop Decisions in New Product Development
Sarangee et al. (2014)	De-Escalation Mechanisms in High-Technology Product Innovation
Eliens et al. (2018)	Rational Versus Intuitive Gatekeeping: Escalation of Commitment in the Front End of NPD
Roeth et al. (2019)	The Interaction of Intuition and Rationality During Escalated NPD Decisions: An Investigation of Decision-Makers' Affective States
Yang et al. (2020)	What Explains Managers’ Escalating Behaviors in a Failing NPD Project? The Impact of Managerial Perceptions of Opportunities and Threats in a Stage-Gate Process

2.6.3 Identified Gaps in the Literature

The relative paucity of research applying behavioural economics and decision-making psychology to high-tech R&D decision-making was evident in the small number (7 of 28, see Table 20) of related articles included in the literature review compared to the relatively high emphasis on traditional economics and analytical methods. The Framework, which critically identifies taking multiple perspectives and distancing strategies to mitigate commitment escalation, relies on this small pool of research for legitimacy. However, the coverage of these mitigations in the high-tech literature was sparse and the inclusion of these terms in the Framework was implied by the inclusion of factors such as gatekeeper independence (Boulding et al., 1997; Eliens et al., 2018; Roeth et al., 2019; Sarangee et al., 2014). It was a small leap to generalise such terms as "multiple perspectives and distancing strategies," but it was nonetheless a leap. In this section, the opportunities for inclusion of new relevant material are expanded on, and this small leap is supported by briefly discussing literature that was not identified in the literature review.

A key focus of this thesis is integrating decision-making from a rational and behavioural sense, and no one has studied this concept more famously than Daniel Kahneman and Amos Tversky. Their work on prospect theory (Kahneman & Tversky, 1979) explores the asymmetry in perception of value that underpins loss aversion and contributes to commitment escalation. It is highly relevant to this thesis. Also highly relevant is Kahneman's book "Thinking, Fast and Slow" (Kahneman, 2011) which addresses framing effects, and the illusion of understanding, providing insights into the challenges in decision-making. However, these concepts are sparsely explored in the context of R&D decision-making, with only Eliens et al. (2018) and Roeth et al. (2019) explicitly exploring rational and intuitive decision-making in the context of NPD.

One powerful way of evaluating multiple perspectives is to involve groups in decision-making. This was well implied in the literature, but rarely stated, and only Schmidt et al. (2001) and Eliens et al. (2018) have explicitly explored this. Anita Woolley and Pranav Gupta have researched many aspects of collective intelligence (Woolley & Gupta 2024) including group decision-making (Woolley et al., 2015). Their work illuminates the conditions for effective group decision-making and underscores the relevance of considering others' perspectives as a way to mitigate bias.

Venture governance is a field with well-established principles for implementing gatekeeper independence. While operating at the level of the firm rather than an R&D project, investors must nonetheless make similar go/no-go decisions at different stages in the firm's evolution. Further, in the case of start-ups, the firm might only have one project/product and start-ups are similar to R&D projects in many ways because they face similar challenges, overcoming technical and market uncertainty. Garima Garg has studied how different governance structures affect decision-making

(Garg, 2020). Her work highlights the importance of aligning governance practices with the venture's developmental stage and strategic goals to effectively manage risks and foster growth. Her approach and objectives seem well aligned to good R&D portfolio management, and should be considered further in the field.

Finally, researchers such as Mark Steyvers and Anoop Kumar have begun exploring AI-assisted decision-making, analysing how AI tools can augment human judgment, particularly in complex environments (Steyvers & Kumar, 2023). Researchers have also begun blending approaches, such as Pranav Gupta and colleagues' work on AI-assisted group decision-making (Gupta et al., 2023). These techniques could augment decision-making by: Providing AI-facilitated guidance through the process, or by having AI provide an opinion on the best option given a blend of subjective and objective criteria and an understanding of the product strategy, or by having AI assess the potential for bias based on a manager's personality and level of investment in the project.

These constantly evolving fields – behavioural economics, collective intelligence, venture governance and AI-assisted decision-making – can powerfully influence the effectiveness of R&D decision-making and present a significant opportunity for future research and refinement of the Framework.

2.6.4 Contribution to the Field

By developing the Framework and highlighting the opportunities for future research, this thesis makes several key contributions, both in theory and practice. It is built on a broad theoretical foundation, incorporating concepts from risk and uncertainty management, real options theory, commitment escalation and bias, and R&D decision-making, and the Framework was synthesised from systematic research based on 28 highly relevant sources identified using a highly-scalable AI-assisted search method. This method adds further credibility to using AI to enhance literature reviews and accelerate and scale the literature search process.

By consolidating the extant literature, this review bridges the research-practice gap and amplifies the practical value of the many academic contributions on which it is based. In doing so, this review also clarified the paucity of behavioural economics and decision making research in the high-tech R&D space.

Designed for high-tech SMEs, the Framework has well-defined applicability, boundaries, and limitations. It consolidates 126 factors into 15 essential factors (groups), reducing complexity and making the myriad factors influencing go/no-go decision-making more comprehensible. It explicitly addresses the risk of commitment escalation by identifying sources of error and advocating a decision-making approach that incorporates distancing strategies and multiple lenses to mitigate bias. By providing a holistic overview of relevant factors, it encourages decision-makers to consider aspects that might otherwise be overlooked, leading to more balanced decisions.

The Framework is process-agnostic and should integrate well with stage-gate and other project governance models, supporting decision-making at critical gates based on structured criteria rather than intuition. By distinguishing antecedents, tools, and criteria, it is deployable because it clarifies when factors apply and how they should be addressed. Antecedents must be considered and implemented in advance, tools may or may not require established data in advance, and criteria must be prioritised and evaluated at the decision point.

2.6.5 Implications for Practice

For researchers – practitioners of literature reviews – the AI-assisted search method provides a template for accelerating and scaling systematic reviews. Further, the high pace of change in the field of AI will continue putting pressure on traditional research methods as organisations seek to leverage the efficiency improvements offered by AI. Researchers would be well advised to heed the call to action, and explore the uses of AI and the many tools that are emerging. As Garry Kasparov, former world chess champion, said: “AI will not replace humans. It will replace those who refuse to work with it.”

Beyond academic contributions, the Framework provides considerable guidance on how to structure and make go/no-go decisions for practitioners. To effectively implement the Framework, managers should self-assess decision-making strengths and weaknesses. Next, they should identify and establish foundational elements such as independent gatekeeping and clear decision criteria. Finally, they can refine decision-making over time by embedding these practices into organisational culture. Immediate improvements may be made by implementing structural changes to the product governance group – that is to say, by getting an independent gatekeeper – and by clarifying decision-making criteria up-front (well before the decision point).

Finally, the approach proposed in the Framework outlines specific guidelines for effective decision-making: (1) Rather than chasing any/all opportunities, it is essential to **have a clear strategy** for the product; (2) Instead of waiting until things are going badly, before negotiating priorities and criteria, decision-makers should **contract options and go/no-go criteria up-front**; (3) Managers should avoid just going with one’s gut or overanalysing by **using a formal decision-making process**; and (4) They shouldn’t let invested parties make portfolio decisions – instead use distancing strategies such as **independent gatekeepers**.

The consolidated research from this literature review suggests that these guidlines, when applied properly, should significantly improve R&D decision-making.

2.6.6 Limitations of the Literature Review

While AI has the potential to significantly accelerate and scale the literature review process, it is not without limitations. This novel and unproven approach carries the risk of introducing errors. To address this, an evaluation of the AI-assisted method is included in Appendix G. Whether AI reduced bias by acting as a second reviewer or introduced bias due to inherent biases remains unknown. However, these effects were expected to be minor relative to human rater errors. AI functioned as a more sophisticated search tool, rather than making final selection decisions or interpreting data. Here, human rater errors refer to this author's ability to maintain focus and reviewer consistency throughout the search without the support of a second reviewer for analytical redundancy. Any biases due to AI were expected to impact quantitative results more than conceptual conclusions, as quantitative outcomes are more sensitive to counting errors.

Repeatability was another unknown of AI-assisted search. AI's interpretation of relevance could vary between runs, introducing potential inconsistencies in article rankings. To assess this, Appendix G provides an empirical evaluation of inter-rater consistency (human vs. AI) and AI self-consistency (repeatability). The results demonstrate that inter-rater agreement and AI repeatability were excellent for this research topic. However, Appendix G constitutes only one small study of the efficacy of AI-assisted search methods, so there remains the potential for errors in the results reported here.

In designing and executing the search, this author adopted a pragmatic approach to defining the search domain. Articles unrelated to decision-making and technology were included if they contributed to answering the research question. Pragmatism was justified because the research sought to explore "How should firms evaluate…?" – which implies a broad range of possible factors – rather than "What is the state of the literature concerning…?", which would require a stricter assessment of the status quo. When decisions arose about whether to exclude an article based on strict criteria or include it due to its relevance to the research question, the latter was preferred. Nonetheless, articles still had to appear in the initial search to be considered, which limited the search domain to technology, decision-making and R&D. This created a tension between search design and execution – the former relied on software and AI, while the latter ultimately required human adjudication. This tension was partially resolved by leveraging AI-assisted search to scale the number of results considerably. However, the search remained focused on decision-making in high-tech R&D, meaning some relevant articles from adjacent fields – such as decision-making bias outside of technology – may have been excluded.

Finally, while developed from credible research, the proposed Go/No-Go Framework has never been tested in the real world. This is an accepted limitation in the scope of this thesis. However, the simulation study in Part B adds credibility to the Framework by demonstrating the value of effective mid-project decision-making, particularly after feasibility studies are completed.

Despite these limitations, this study provides a strong foundation for both theoretical exploration and practical application, as outlined in the concluding section.

2.7 Conclusion of Literature Review

Part A sought to integrate the extant literature regarding go/no-go decision-making in high-tech R&D. A systematic literature review was conducted, and a systematic approach was used to produce a novel decision-making framework focused on go/no-go decisions in new product development. This directly addressed a gap in the literature and the research question by proposing a Framework that can be used to support go/no-go decisions and improve process maturity.

This thesis makes several key contributions to both research and practice by integrating a broad range of management theories pertinent to R&D, and advances academic understanding of high-tech R&D governance. At the same time, the proposed Go/No-Go Framework provides a practical tool for decision-makers.

For researchers of R&D decision-making, the review serves as a snapshot of the state of the literature and signposts some paths forward. The AI-assisted approach offers a template for leveraging AI in research without compromising ethical values. For R&D project managers seeking to translate uncertainty into manageable risk, value is added by providing a framework for managing project risks related to decision-making. Project managers can support portfolio managers by running effective gate reviews and surfacing factors that may lead to poor decisions. For portfolio managers seeking to apply resources effectively, the Framework provides a decision-making checklist and guidance on structuring project governance. In particular, the independence of gatekeepers is emphasised, as is the need for formality in the process, tight contracting around no-go conditions and strategic alignment. Most importantly, the literature and the Framework warn of commitment escalation and the biases that can lead to it.

For practitioners, the Go/No-Go Framework provides a structured approach to improving decision-making. By clarifying decision criteria, ensuring independent gatekeeping, and adopting distancing strategies, firms can mitigate escalation bias and enhance project governance.

While this research provides a structured framework for decision-making, several areas require further investigation. Future studies could explore empirical validation of the Framework, deeper integration of collective intelligence in decision-making, and evolving applications of AI-assisted governance models.

By integrating insights from behavioral economics, governance structures, and AI-assisted decision-making, the Framework provides a foundation for future empirical research and practical application in high-tech R&D governance. Part B continues this work by developing and presenting an illustrative simulation of decision-making approaches and begins with the rationale for simulation over real-world empirical studies.

3 Part B: Simulating R&D Investment Decisions – An Agent-Based Approach

The literature review identified many factors relevant to R&D decision-making. While the outcomes from the literature review have theoretical value, it is pertinent to seek validation of the Framework. However, as discussed in the introduction, strong empirical validation is difficult, if not impossible. One significant challenge with this research topic is a scant amount of reporting. At the firm level, good data is relatively easy to find. Markets, government agencies, and venture capitalists routinely analyse and report on firms' birth, life, and death (U.S. Bureau of Labor Statistics, 2024). The same is untrue for internal R&D projects with negligible reporting requirements, particularly in the private sector. Another challenge is repeatability. In comparison studies, the environment and capabilities within teams would differ, or performed by the same team, learning and path dependency would perturb the outcomes. Using large samples and statistical techniques could overcome these challenges but is not practical as this approach would require large numbers of teams and even larger budgets.

Thus, simulations are common among R&D decision-making researchers, and most are computerised. The most widespread technique is Monte Carlo analysis. Medaglia et al. (2007) used it to test an evolutionary method for project selection, Choungsirakulwit and Sutivong (2007) to test a quantitative model for balancing and optimising a portfolio of R&D projects and Ayala-Cruz (2016) to test an enhanced project risk management framework. Cooper et al. (1997) report that Proctor & Gamble uses Monte Carlo to support project selection, and similarly, Chan et al. (2007) report that Merck uses it for project risk evaluation. Simulation offers a robust, repeatable and inexpensive environment for evaluating the efficacy of the proposed Framework. For these reasons, the decision was to test the Framework using simulation.

This simulation study aims to:

1. Validate the Framework by illustrating the application of the proposed decision-making approach compared to other approaches, and
2. Test the sensitivity of the outcomes to varying decision-making strategies and feasibility phase investment levels.

This literature review in Part A added to the body of knowledge by developing a framework that focuses on go/no-go decisions in an NPD process, and the simulation will add value by providing empirical support for the Framework, highlighting the impact of different decision-making strategies. The results will be relevant to academic studies and practical approaches to R&D decision-making. Whereas other studies (Farshchian & Heravi, 2018; Kettunen & Salo, 2017; Saiz et al., 2024) often focus on one particular approach, this work will compare multiple decision-making approaches in different environments to inform the applicability of the Framework.

3.1 Simulation Approach

Building on the theoretical foundation, the simulation approach aimed to reflect objectivity in decision-making. The literature review highlighted objectivity as a desirable attribute for an R&D decision-maker, so a positivist philosophy and empirical approach were adopted. This alignment supported the simulation's aims for two reasons: first, a highly repeatable study was required to compare different decision-making strategies; second, an empirical approach added rigour and philosophical balance to the inductive reasoning and pragmatism of the literature review.

The process began with analysing the Framework to determine which factors were amenable to simulation and which were to be excluded. Factors that could be quantified or represented algorithmically – such as investment thresholds and decision criteria – were included. In contrast, elements such as culture or politics, which were difficult to parameterise, were excluded. An agent-based approach was used in which the "performers of the activities [were] active agents" (Sulis & Taveter, 2022, p. 7).

Three agents were constructed, representing three decision-making models – ROI-based, risk-based, and gated – and modelled in Google App Scripts (Google, 2024), incorporating ROI, risk level, seed funding and feasibility investment into the decision criteria. Two of the three agents were chosen because they represent the most common decision-making strategies (ROI-based and risk-based) (Boulding et al., 1997; Chan et al., 2007; Cooper et al., 2000; Hauser & Zettelmeyer, 1997; Jagle, 1999; MacMillan & McGrath, 2002; Neely III & De Neufville, 2001; Nielsen et al., 2024; Perlitz et al., 1999; Pillai et al., 2002; Sheasley, 2000; Smit & Trigeorgis, 2007; Vaculik et al., 2019), and the third (gated) was chosen to represent the introduction of mid-project decision gates.

A discrete event simulation was created to model the agents' actions following guidance from Zeigler et al. (2018) in "Theory of modelling and simulation: Discrete event & iterative system computational foundations (3rd ed.)". Discrete event simulation involves identifying specific events (e.g. decision points) and skipping the intervening time interval so that computational resources are not wasted, and the simulation runs quickly and efficiently.

Further, a Monte Carlo approach was taken (Brandimarte, 2014), which involved creating a portfolio of 1000 repeatable projects from scaled beta distributions, each with ex-ante (perceived) and ex-post (actual) cost and return. Beta distributions were used because they can be shaped to model the cost and revenue distributions typical in R&D. Monte Carlo simulation is a commonly used approach to incorporate uncertainty by running simulations many times with different parameters and control variables (Kazak & Pohlmeier, 2019; Kettunen & Salo, 2017; Legkokonets et al., 2020; Narayanan et al., 2023; Saiz et al., 2024; Xuan et al., 2021). The simulation goal was always to *maximise profit over the portfolio*. Each agent evaluated the portfolio based on perceived inputs, and the outcomes were aggregated.

To test the agents under high uncertainty and to negate the possibility of artificially favourable conditions with the beta distributions, the three agents evaluated a second portfolio of 1000 projects constructed from a uniform distribution with the same goal. Uniform distributions have no mode, so the cost and return for a project can be any value in their range with equal probability. This approach was difficult to validate against the literature because other similar articles that empirically explore project selection either didn't specify the underlying distribution or used only one distribution (Aguilar-Rivera & Valenzuela-Rendón, 2017; Conversano & Lizzeri, 2012; Farshchian & Heravi, 2018; Kazak & Pohlmeier, 2019; Kettunen & Salo, 2017; Narayanan et al., 2023; Saiz et al., 2024; Teglio et al., 2009; Xuan et al., 2021). This use of two types of distribution (beta and uniform) positioned the agents in a semi-realistic and then hostile environment and thus bracketed the likely operating environment of decision-makers.

Table 21

Search Parameters For AI-Assisted Search Adoption

Scopus Keywords	simulation AND (monte carlo OR agent-based) AND (stage-gate OR stage gate OR decision gate OR project selection)
Earliest Year	2010
Journals	Any
Citescore	Any
AI Review Keywords	simulation experiment, agent-based model, monte carlo simulation, decision strategy, portfolio performance
AI Research Question	How have simulation methods been used to study and compare decision-making strategies in R&D contexts?
AI Relevance Cutoff	$AI\_R \geq 6.0$

3.1.1 Confirmation of Simulation Approach: A Brief Literature Review

To inform and validate the simulation approach, a brief literature search was conducted using the AI-assisted method described in the literature review, with the search parameters in Table 21 above. The search was designed to identify articles that inform the simulation approach and focus on simulations relevant to R&D decision-making either at initial project selection or mid-project at a decision gate.

The search results (Table 22) identified 22 articles from an initial pool of 565 abstracts. The identified articles are organised on a grid with the predominant simulation method as the rows and whether or not the article compares decision-making methods as columns.

Table 22

Simulation Approaches from Literature

	Does the article compare decision-making methods?	
Method	No – single decision-making approach	Yes – multiple decision-making approaches
Agent-based	Ilie-Zudor & Monostori (2009), Li & Zhang (2020), Liu & Louis (2016), Singh et al. (2022) [4 articles]	Li et al. (2012) [1 article]
Monte Carlo	Bard & Feinberg (1989), Kettunen & Salo (2017), Kholany & Abdelsalam (2017), Marcondes (2021), Marcondes & Marcondes (2024), Mavrotas & Makryvelios (2021), Mavrotas & Pechak (2013), Panadero et al. (2020), Pendharkar (2014), Saiz et al. (2024), Tinoco et al. (2018) [11 articles]	Barucke Marcondes et al. (2016), Marcondes & Vilela (2022), Medaglia et al. (2007), Naldi et al. (2019), Shakhsi-Niaei et al. (2011), Simplício et al. (2012) [6 articles]

These results show that Monte Carlo and agent-based simulation have been commonly used by other researchers and also that comparing different decision-making methods is not uncommon, although many different decision-making methods are explored: Barucke Marcondes et al. (2016) test the Mean-Gini approach against stochastic dominance. Li et al. (2012) compare suboptimisation, waiting, and utility strategies for conflict resolution; Marcondes and Vilela (2022) compare four different multi-criteria decision-making (MCDM) strategies; Medaglia et al. (2007) explore an evolutionary method versus stochastic parameter space investigation; Naldi et al. (2019) compare profit-focused and fairness-focused budget allocation methods; Shakhsi-Niaei et al. (2011) compare a proposed framework for project selection under uncertainty against a deterministic approach; and Simplício et al. (2012) investigate Markowitz's optimisation against profitability index ranking. Generally, two to four decision-making approaches are compared, but there is no consensus on which approaches should be used. These diverse comparison studies and the commonality of Monte Carlo and agent-based approaches support the simulation approach described here and the selection of three agents.

The 22 articles in Table 22 above were further examined to determine:

1. Whether the article focused on initial project selection or mid-project decision gates.
2. How uncertainty was modelled in the simulations, particularly whether or not multiple probability distributions had been used to construct portfolios of projects.

To the first point, it was found that no (zero) articles modelled mid-project decision gates. Within the scope of the brief literature search, including a decision gate mid-project in simulation study is a novel approach.

To the second point, most of the 22 papers did not disclose the statistical distribution underlying model parameter selection, and the remainder used only a single distribution – most commonly normal or uniform – and did not generate alternative portfolios of projects. That is to say, they tested their selection method, perhaps against other methods, but never against different simulated environments. This leaves the approach of comparing behaviour across two distributions without the legitimacy of being among academic peers. However, a comforting argument can be made that the approach taken here is over and above that taken by the 22 selected articles.

This brief review adds legitimacy to the simulation approach by confirming the popularity of Monte Carlo and agent-based methods in R&D simulations. It also highlights the novelty of simulating mid-project decision gates and using multiple probability distributions to test agent performance in different environments. The next section develops the detailed simulation methodology based on the above approach.

3.2 Simulation Method

3.2.1 Factor Selection and Model Development

As indicated in the approach, it was logical to begin by analysing the Framework to determine what could and should be simulated. Generally, subjective factors were harder to encode than objective ones, though some objective factors also posed challenges. Factors like opportunity cost and resource constraints would have added complexity without necessarily affecting result validity (assuming consistent conditions across runs). Further, implementing agents employing simulated simulation or benchmarking[3] implied a broader decision perspective and more complex coding than simple criteria evaluation. Other multi-criteria and ranking tools required subjective judgment and were not easily translatable into code. In summary, factors were excluded primarily for two reasons: (1) They added complexity or implementation difficulty, and (2) they were subjective and would have required AI to implement, further increasing complexity. Thus, the factors in Table 23 were excluded from the simulation model.

Table 23

Go/No-Go Framework Factors Excluded From the Simulation

Type	Factor	Rationale for Exclusion
Antecedents	Capabilities – customer orientation, market awareness	Subjective and difficult to encode
Criteria	Strategy – balanced portfolio	Subjective and difficult to encode
	Finance - opportunity cost, NPV	Simplicity
	Resources – capacity, management support, capability	Simplicity
Tools	Planning & communication – roadmaps, visual indicators	Subjective and difficult to encode
	Option analysis – benchmarking, sensitivity analysis, simulation, resource-capacity analysis, decision tree	Subjective and difficult to encode
	Evaluation – scoring model, selection matrix, force-ranking	Subjective and difficult to encode
Sources of Error	Bias – impression management, commitment bias, self-justification	To demonstrate the effect of rational decision-making
	Mismanagement – too many projects, misuse of NPV, gates with no teeth	Subjective and difficult to encode

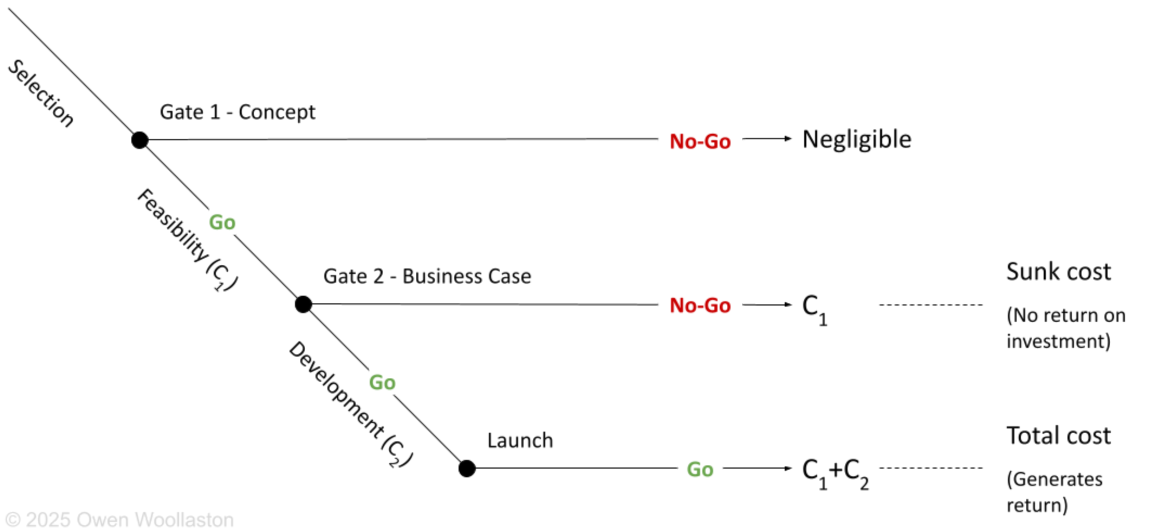
In real-world R&D, project selection follows structured decision gates to manage risk and ensure alignment with strategy, and the Framework was designed to operate within a stage-gate process or other portfolio management system in which decision gates are used. To mirror this, a simple stage-gate process was included in the simulation model. A typical stage-gate process includes at least four phases: Selection, feasibility, development and commercialisation (Choungsirakulwit & Sutivong, 2007). Commercialisation was ignored here for simplicity, so the model included two gates between the first three phases in this simulation. This structure enabled the simulation model to reflect real-life R&D decision-making more closely.

[3] That is, simulating an agent using simulation or benchmarking.

Following the guidance from Zeigler et al. (2018), the simulation model was formulated in terms of a stage-gate process and then converted into a discrete event model. A discrete event simulation is a computer simulation in which only events of interest are modelled, and the intervening time is skipped. As Zeigler puts it: *"The prerequisite of discrete event modeling, therefore, is to have a means to determine when interesting things happen … Events can be caused by the environment … or … can be internal events and the [model] itself determines their time of occurrence"* (Zeigler et al., 2018, p. 84). In this simulation, the events were easily identified as the gates in the stage-gate process, as shown in Figure 20.

Figure 20

Discrete Event Model Formulation



The first event – the "concept" gate – followed the selection phase and served as the initial screening of projects to decide whether to invest in a project or pass and assess the next one. The second "business case" gate followed feasibility and served as the more thorough business decision to commit to the full development. In the selection phase, ideas are assumed to be unlimited and costless; therefore, costs only accrued during feasibility and development. Since the feasibility phase aimed to raise confidence in the project's viability, improved forecast data was a key outcome of the phase. Returns were realised at the end of development with the unconditional release of the product ("Launch") and represented the NPV over the lifetime of the resultant product. It was, therefore, (1) necessary to make a project selection decision at the concept gate and (2) possible to make a go/no-go decision after feasibility at the business case gate.

Table 24

Go/No-Go Framework Factors Included in Simulation

Category	Factors	Modelling Approach
Approach	A formal process that ensures managers make strategically aligned decisions by evaluating options objectively using distancing strategies and multiple lenses.	Include a second go/no-go gate between the feasibility and development phases Decide objectively based on criteria
Antecedents	A defined strategy	Test different decision-making approaches
	Capabilities – organisational agility, project tracking	Encode tracking in portfolio accounting Include a second go/no-go gate between the feasibility and development phases
	Contracting – clear requirements, budget allocation, precommitted rule	Encode goals and seed funding into the simulation
	Oversight – project manager, gatekeeper independence	Encode objectivity into project selection and go/no-go criteria
Criteria	Strategy – alignment, outcomes	Set the goal of portfolio profit maximisation Test different decision-making approaches
	Finance – costs, budget, benefits, market value	Test ROI as a primary decision-making approach
	Unknowns – uncertainty, risks, assumption validity, technical feasibility	Test risk as a primary decision-making approach Encode uncertainty into cost and returns
Sources of Error	Data quality – inadequate data, estimation error	Encode uncertainty into cost and returns
	Bias – impression management, commitment bias, self-justification	Different agents implement differing levels of commitment at gates
	Change – market volatility	Implicit as unknowns in cost and returns

With the model structure established, the next step was to identify the decision-making factors to include in the model. These are summarised in Table 24, along with the modelling approach for each key factor. Factors from the Framework that require human interaction or represent human errors were omitted, and encodable factors are listed. Specifically, agents made decisions based on perceived ROI and risk and were constrained by budget and initial seed funding. Further, they had to make an initial decision at the concept gate and (depending on the agent) could choose to abandon the development at the business case gate (based on better information). For a realistic simulation, information at both gates was imperfect (Pace, 2013), so risk and uncertainty were encoded into the model. As Schwartz and Moon (2000) identify, there are three underlying uncertainties in product development: Investment cost, future payoffs, and the possibility of failing before the project is completed. Probability distributions were used to model development costs and returns to capture these points. The information available ex-ante was taken from selected parameters of the distributions. In contrast, the actual cost and returns were taken as randomly selected values from the distributions. A failure occurred when the actual cost bankrupted the portfolio, and with no more funds, the simulation ended.

Finally, the goal was to maximise the value of the portfolio. In the real world, this might be a naive and short-sighted goal, and typically, portfolio managers try to strike a balance between market share, portfolio balance, and immediate and future returns. However, the approach was acceptable in the simulated environment because it enabled a like-for-like comparison between different decision-making approaches.

3.2.2 Simulation Design

With the model structure defined and decision factors selected, implementation became the next step. To bring the simulation to life, Google Sheets and App Scripts (JavaScript) were used to construct 1000 simulated projects, each described by a pair of beta distributions: one for the project cost and one for the revenue. The Python-based library SimPy (SimPy, 2024) would have been a natural choice for this type of simulation, but it would have required more direct programming and lacked the natural integration of Google Sheets plus App Scripts, which allowed the use of all of Google Sheets analytical and statistical functions as well as graphing tools without additional coding. However, the choice was largely a matter of convenience because the visualisation tools available in Google Sheets were easier to use, and JavaScript was very familiar to this author. (Google, 2024).

The decision-making process is captured in Figure 20 (page 87). In the selection phase, 1000 projects were constructed. At the concept gate, a simulated agent ("agent") decided whether to commit to the first project or pass and evaluate the next project, and so on. Three agent types were modelled to simulate different approaches to project selection. These agents represent common decision-making strategies observed in corporate R&D: (1) a purely financial perspective (ROI Agent), (2) a risk-aware approach (Risk-Aware Agent), and (3) a more adaptive strategy incorporating staged decision-making (Gated Agent). Projects were always evaluated in the same order, and random numbers were generated only once so that projects remained the same from run to run for repeatability. Agents could not look ahead at the next project(s) to decide which was better, and they had to decide based only on the information presented regarding the current project. Three agents were constructed, each with a different decision-making approach:

1. **ROI Agent** – The agent was only aware of the expected cost and revenue at the concept gate and committed to the entire project, i.e. feasibility ran directly into development.
2. **Risk-Aware Agent** – As above, with one additional criterion: The risk was also evaluated at the concept gate. The coefficient of variation was used as a proxy for normalised risk and is defined below.
3. **Gated Agent** – As mentioned above, the decision to abandon or continue at the business case gate was made based on improved information about the project cost. Risk was ignored and only ROI was used. Revenue information remained unchanged and retained the same variability as at the concept gate.

Table 25 defines the gate criteria and the agents they applied to, and the software details are in Appendix E (simulation platform) and Appendix F (agents). All agents were constrained by a funding limit and had to evaluate the same set of projects. If the agent passed on a project, no cost was incurred, but if it committed at the concept gate, then feasibility costs were incurred, and the possibility of a loss arose if the resulting revenue was low.

A critical feature of this design was that the ROI Agent and Risk-Aware Agent model commitment escalation by fully committing to the whole development at the outset. This mirrors scenarios in which managers are ignorant of costs or risks, or are sufficiently biased to continue the development despite indications to the contrary. Conversely, the Gated Agent committed in tranches, first with the feasibility phase, then the development – behaviour equivalent to objective decision-making following pre-committed rules.

Table 25

Simulation Decision Criteria

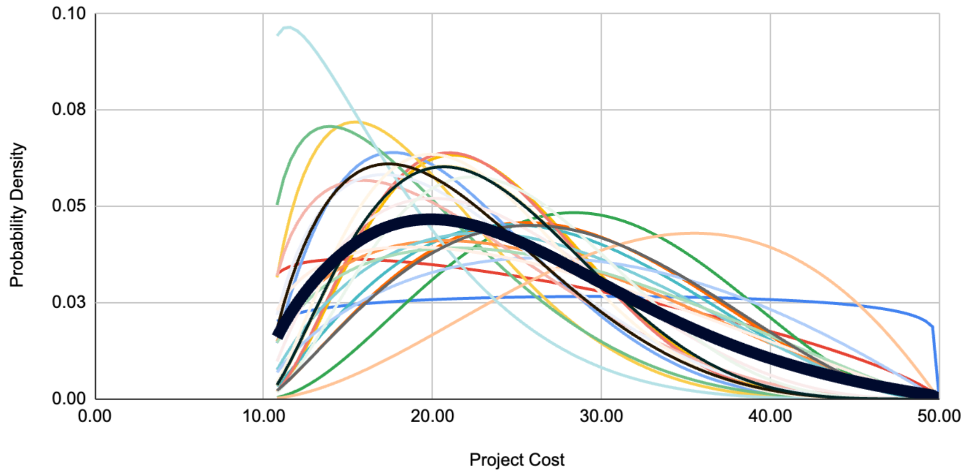
Decision Criterium		Description	Agent/s
Funding	$EC_i = Mc_i$ $F_i > EC_i$	Initially, the expected cost (EC_i) equals the mode of the cost distribution. At all gates, the available funds must exceed the expected project cost	All
ROI	$ER_i > (1 + ROI_{th}) \bullet EC_i$	At all gates, the expected ROI (ER_i) must exceed the threshold, ROI_{th}.	All
Risk	$CV_i < CV_{th}$	At the concept gate, the perceived risk CV_i must be less than the risk tolerance threshold CV_{th}	Risk-Aware
Gate	$EC_i = (Mc_i + G(AC_i - Mc_i))$ $ER_i > (1 + ROI_{th}) \bullet EC_i$	At the business case gate, the expected cost (EC_i) is closer to the actual cost by G (percent), then the ROI criterion applies.	Gated

Where:

$i \subset [1, N],\ N = 1000$	Simulation run index and the total number of simulations
$G = 25\%\ or\ 50\%$	Percentage improvement in cost accuracy at business case gate.
$C_i = Beta(\alpha c_i, \beta c_i, lc_i, hc_i)$	Four-parameter (scaled) random beta distribution describing the project cost at run *i*, with mode Mc_i, and standard deviation σc_i.
$CV_i = \sigma c_i / \mu c_i$	Coefficient of variation of C_i, as a proxy for normalised risk in the project cost
$R_i = Beta(\alpha r_i, \beta r_i, lr_i, hr_i)$	Four-parameter (scaled) random beta distribution describing the revenue at run *i*, with mode Mr_i, and standard deviation σr_i.
AC_i, AR_i	Actual (ex-post, sampled) cost and revenue at run *i*, respectively
$F_0 = F_{seed}$ If the project is executed, $F_{i+1} = F_i + (AR_i - AC_i)$ Otherwise, $F_{i+1} = F_i$	Beginning with seed funding, F_{seed}, F_i is the available funds at sample *i*, which is the cumulative profit (or loss) of previously executed projects, and F_N is the final value of the portfolio after all projects are evaluated or executed.

Figure 21

Sample of Cost Distributions

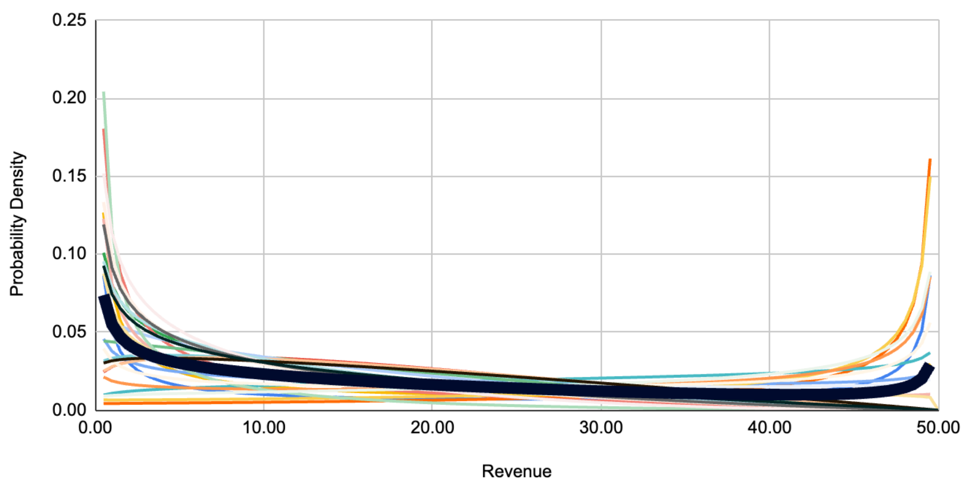


The cost of each project was defined by a beta distribution with randomly generated shape parameters such that $\alpha c \in [1.0, 2.0]$ and $\beta c \in [1.0, 5.0]$ and scaled on the range $Xc \in [10.0, 50.0]$. This produced the characteristic right-skewed (log-normal) shape typical of distributions of effort in new product development, as pictured in Figure 21 (McConnell, 2006). The bold line represents the average of the distributions. Note that not all of the generated distributions were right-skewed. This was deliberate and represents an occasional project that is more costly than expected or does not follow the typical pattern of costs. The beta distribution was chosen for its ability to flexibly define a diverse range of distribution shapes without changing the underlying mathematics.

Similarly, the return from each project was defined by a beta distribution with randomly generated shape parameters such that $\alpha r \in [0.2, 1.0]$ and $\beta r \in [0.2, 3.0]$ and scaled on the range $Xr \in [0.0, 50.0]$. This produced the characteristic bathtub curve associated with product revenues (Anderson, 2006; Abyad, 2020) – many failing early, but some highly successful – pictured in Figure 22. Note the adverse conditions where the cost must exceed ten units, but revenue can be zero. This is more realistic and ensures that an agent is unlikely to profit by making random choices.

Figure 22

Sample of Profit Distributions

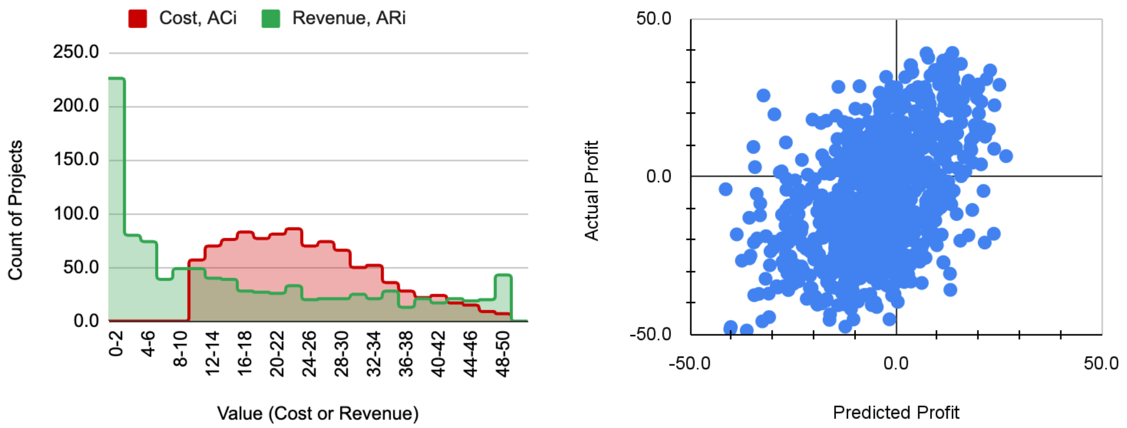


For the Risk-Aware Agent, the coefficient of variation – CV_i, defined as the ratio of the standard deviation to the mean – was used as a proxy for normalised risk. This was useful because it produced a dimensionless number that could be used to compare risk across portfolio runs and was valid because it met the statistical requirement for use, which was that the distributions have a clearly defined zero point – in this case, zero cost or revenue. (Everitt, 2010)

Ex-ante, the agent received information about the distributions as pictured above, but ex-post, the actual values were known. For the 1000 projects (the "Beta Portfolio"), the actual cost and revenue are pictured in Figure 23. From this, it is easy to see that costs exceeded revenue in most cases because the cost (red) is weighted to the right compared to the revenue (green).

Figure 23

The Beta Portfolio



The initial estimates (as perceived by the agents) for cost and revenue were selected based on the mode of the underlying distributions. The distribution's mode was chosen over the median or mean because of the cognitive tendency to recall information more frequently observed prior, a bias called the "availability heuristic" (Harvey, 2020; Tversky & Kahneman, 1973). Think of the travel time between two frequented locations. The reader will likely recall the most frequently observed travel time (the mode, like the peaks in Figure 21 on page 91), not the actual mean or median, which requires mathematics (and thus conscious processing) to obtain. This tendency occurs in estimation where the value managers expect is the one they have observed most often.

With the model, agents and project portfolio defined, the next section considers the modelling assumptions and their legitimacy.

3.2.3 Assumptions

Several simplifying assumptions were made to maintain a controlled and repeatable simulation environment and reduce the model complexity to fit within the scope of this thesis. While these assumptions abstract certain real-world complexities, they allowed for an efficient and focused comparison of decision-making strategies without confounding variables.

Generally, abstraction was not a concern as long as the illustrative nature of the simulation was preserved as a tool for comparing different decision-making approaches. Assumptions in this category were:

1. There were no overheads, and reinvestment (R&D intensity) was 100%.
2. Abandoned projects had no residual value.
3. All revenues were accrued at release (end of development); that is, AR_i was the net present value of all future revenues at release.
4. Objectivity was assumed in the mathematical implementation of the decision criteria.
5. The environment did not change during the simulation or from run to run.
6. Projects were independent, resourcing was unlimited, and there were no resource conflicts.
7. Project cost and returns, while unknown, were constant during a simulation run.

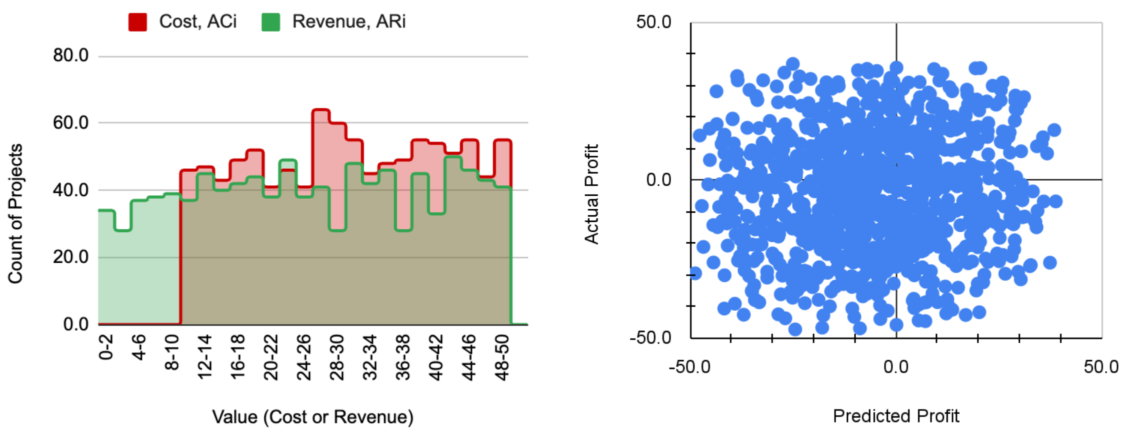
More significantly, it was assumed that using a beta distribution was appropriate. This was an assumption of convenience and simplicity, but the validity of the approach remains uncertain. There is some support from the literature: The cost (effort) distribution in software developments has been shown to follow a log-normal distribution (McConnell, 2006), which is similarly shaped to the cost distributions here with the exception that the log-normal distribution is always a right-skewed bell curve and, therefore, cannot model projects that do not conform to the pattern of costs. Also, log-normal distributions have no upper bound (Everitt, 2010). Further, the use of beta distributions to model the development effort has been established for many years, beginning with the PERT distribution (Malcolm et al., 1959) and its application to US Navy projects. With their unbounded right tail, log-normal distributions might be more technically correct but could not be used to generate exceptions to the rule; therefore, the beta distribution was selected.

Product revenues are known to follow various distributions (Anderson, 2006; Abyad, 2020), so it is not possible to identify a “correct” one. Another approach to valuing R&D projects is to assume that “the gross value of a project follows a random walk (a stochastic process) and using a tree structure to depict uncertain movement over time” (Choungsirakulwit & Sutivong, 2007), but this requires a different (time-stepped) simulation approach and is not compatible with discrete event simulation. The beta distribution was also selected for convenience for product revenues because it could be configured to produce the characteristic bathtub shape common in product development (Anderson, 2006; Abyad, 2020).

To avoid depending solely on the beta distribution, the impact of an alternative distribution was tested by generating a second portfolio of 1000 projects using a uniform distribution (the "Uniform Portfolio") – represented in Figure 24. For the uniform distribution, the mode can be any value (which is also appropriate for this simulation), so a random value for cost and revenue was chosen from the distribution range. This means the predicted cost and revenue did not correlate to the actual cost and revenue, and the projects were equivalent to managing in extreme uncertainty. Using the Uniform Portfolio, all simulations were repeated for each agent – ROI, Risk-Aware and Gated.

Figure 24

The Uniform Portfolio



3.2.4 Verification of Model Implementation

The model was implemented in JavaScript and, as with any software development, required testing to verify it functioned as intended. In this case, the procedure was exceptionally simple. Each agent was run across a portfolio of projects, and the first 10 results were manually inspected (sense-checked). Then, the expected agent behaviour was modelled independently in a spreadsheet and, once again, compared with the JavaScript output. The process was repeated for the Beta Portfolio and Uniform Portfolio to ensure consistent behaviour. After initial debugging, the agents were all found to perform as expected (as defined in the method above), and the model implementation was verified.

In summary, the simulation model combined discrete event and Monte Carlo methods to evaluate three decision-making strategies across two portfolio environments. This setup enabled a controlled, repeatable comparison that reflected both structured and uncertain project conditions. The following section presents the simulation results and compares agent performance.

3.3 Simulation Results

Having designed and implemented the simulation model, it was run for each scenario – 3 agents times 2 portfolios. The results are presented agent-by-agent in the following sections, comparing their performance between the Beta Portfolio and Uniform Portfolio.

3.3.1 ROI Agent Results

The first results came from the ROI Agent evaluating the Beta Portfolio, where it made simple decisions based only on perceived ROI and committed to the whole project at the concept gate. The results from the Beta Portfolio and Uniform Portfolio simulations for the ROI Agent are shown in Table 26 and Table 27 below, respectively. The tables show portfolio returns after evaluating or executing all 1000 projects (depending on decision criteria), equal to F_{1000}. The colours indicate a heatmap of losses (red) and gains (green), and the results do not include the initial seed funding (thus the losses). Empty grey cells indicate that no projects met the initial investment criteria (at the concept gate), and the entire portfolio was passed over.

Table 26

ROI Agent (Beta Portfolio)

		Seed Funding, F_{seed}					
		20	**30**	**50**	**100**	**200**	**500**
Expected ROI, ex-ante, ROI_{th}	**-50%**		-25	-25	-85	-175	-494
	-40%		-25	-25	-85	-175	-489
	-30%		-25	-25	-82	-168	-460
	-20%		-25	-25	-93	-181	-10
	-10%		-25	-39	-71	-192	734
	0%		-25	-46	1246	1246	1246
	20%		-25	-46	1686	1686	1686
	40%		-25	-46	1600	1600	1600
	60%		-25	-46	1356	1356	1356
	80%		-21	1083	1083	1083	1083
	100%		-21	759	759	759	759
	150%		-21	230	230	230	230

Losses were limited by the seed funding (F_{seed}), and path dependency had a role to play; that is to say, if the initial projects were unlucky and the seed funding was exhausted early, the subsequent (possibly lucrative) projects were never evaluated. This path dependency is realistic and reflects the observation that many small ventures fail due to cash flow problems. With the Beta Portfolio, a minimum seed funding of 50 was required to make a profit, and this required a bearish approach (ROI_{th} > 80%). Conversely, given seeding funding of 500, a bullish approach with ROI_{th} > 0% made a good return. This highlights the advantage of generally larger, well-funded firms over less well-funded startups and SMEs: If the seed funding was adequate and the ROI threshold (ROI_{th}) was high enough, the portfolio made a profit.[4] However, being too conservative (ROI_{th} > 100%) limited the portfolio's value by excluding profitable projects.

With the Uniform Portfolio (Table 27), using ROI to predict outcomes was impossible. This was because the Uniform Portfolio did not correlate the predicted return with the actual return (R^2 = 0.002), so ROI had no predictive value, and all approaches failed regardless of seed funding.

Table 27
ROI Agent (Uniform Portfolio)

		Seed Funding, F_{seed}					
		20	**30**	**50**	**100**	**200**	**500**
	-50%		-5	-29	-67	-172	-454
	-40%		-5	-14	-52	-165	-462
	-30%		-5	-14	-52	-184	-476
	-20%		-5	-25	-62	-183	-475
	-10%		-5	-25	-71	-177	-471
Expected ROI, ex-ante, ROI_{th}	**0%**		-5	-25	-61	-152	-474
	20%		-5	-25	-82	-175	-480
	40%		-5	-9	-77	-188	-479
	60%		-4	-32	-64	-169	-483
	80%		-4	-27	-67	-169	-468
	100%		-4	-27	-73	-176	-473
	150%		-4	-34	-75	-187	-206

In this unpredictable environment, limiting seed funding was an advantage (because it limited losses), and seed funding of F_{seed} = 50 was adopted for the remainder of the simulations because a relatively small seed funding is more representative of SME finances. Seed funding of 50 represented sufficient initial funds to invest in one or two projects.

[4] This result is beside the point of the thesis, which focuses on funding-constrained SMEs. It also overlooks the cost of the working capital (keeping a larger seed fund in reserve) by assuming that large firms do not stretch their investment funds across many portfolios. Thus, from an individual portfolio perspective, large firms can also suffer from seed funding constraints, depending on their operating model. The difference is that failure of the portfolio does not mean death of the firm.

3.3.2 Risk-Aware Agent Results

Unlike the ROI Agent, the Risk-Aware Agent had an additional criterion to consider before investing (again in the whole project): The coefficient of variation (CV_i , the "risk") was used as a measure of uncertainty and enabled an extra dimension of selectivity. *CV* values ranged from 0.25 to 0.41 (dimensionless ratio of standard deviation to mean). Table 28 shows the results for the Beta Portfolio as evaluated by the Risk-Aware Agent with seed funding 50. The agent chose projects for which the risk is lower than CV_{th} and the expected ROI greater than the ROI threshold, ROI_{th}. Notable observations from the results are:

1. The pink region (top right) is where the risk was too high and the expected ROI too low, resulting in portfolio failure.
2. The highest returns (darkest green) tended to encircle the pink failure region, creating a drop-off ("cliff") between success and failure and a region of high sensitivity to selection criteria.
3. Being too conservative with risk (CV_{th} < 0.28) excluded all projects, as did being too conservative with ROI.
4. As expected, the highest risk (right-hand column) matched the corresponding (F_{seed} = 50) column for the ROI Agent above. This is because the simulation became indiscriminate to risk at that point and reverted to simple ROI-based decision-making.

Table 28
Risk-Aware Agent (Beta Portfolio, F_{seed} = 50)

Expected ROI, ex-ante, ROI_{th}	Risk Threshold, CV_{th}						
	0.25	**0.28**	**0.30**	**0.33**	**0.36**	**0.38**	**0.41**
0%			16	571	-41	-46	-46
10%			16	540	-41	-46	-46
20%			15	666	-41	-46	-46
40%			-14	510	1191	-46	-46
60%			5	381	972	-46	-46
80%				199	678	879	1083
100%				58	326	555	759
120%				13	272	414	577
140%				13	173	234	359
160%					125	119	217
180%					73	52	126
200%					9	9	82
250%							

With the Uniform Portfolio (Table 29), the same underlying randomness that defeated the ROI Agent above (unpredictability of returns) caused the measure of risk (*CV_i*) to be useless as a selection measure, and, as for the ROI Agent, all approaches fail. That is, all except three, which are explained by simple luck.

Table 29

Risk-Aware Agent (Uniform Portfolio, F_{seed} = 50)

		Risk Threshold, CV_{th}						
		0.23	**0.39**	**0.54**	**0.70**	**0.85**	**1.01**	**1.16**
Expected ROI, ex-ante, ROI_{th}	**0%**		-28	-31	-17	-8	-25	-25
	10%		-28	-6	-17	-8	-25	-25
	20%		124	-4	-17	-8	-25	-25
	40%		-21	-40	-12	-30	-9	-9
	60%			-4	-28	-28	-32	-32
	80%			25	-14	-35	-27	-27
	100%			10	-31	-35	-27	-27
	120%				-32	-4	-8	-8
	140%				-4	-4	-38	-38
	160%				-34	-38	-34	-34
	180%				-27	-23	-32	-20
	200%					-40	-45	-28
	250%						-17	-26

3.3.3 Gated Agent Results

Unlike the ROI Agent and Risk-Aware Agent, which committed fully at the concept gate, the Gated Agent had the option to abandon the project after the feasibility phase (at the business case gate). It made the initial investment decision using ROI only, identical to the ROI Agent. However, at the business case gate, it received a refined cost estimate, recalculated the ROI, and reassessed the go/no-go decision (without altering its risk perception). Risk was not considered at either gate.

Two simulations were run to test portfolio profit sensitivity to improved cost estimates: one with a 25% improvement (bringing the estimate a quarter of the way toward actual cost) and another with a 50% improvement (halfway to actual cost).

3.3.3.1 Gated Agent Results – 25% Improvement

For a 25% improvement in the cost estimate, Table 30 presents the Gated Agent's performance on the Beta Portfolio, using the same seed funding of F_{seed} = 50 and no risk threshold. The agent made two key decisions: (1) At the concept gate, decide whether to invest a percentage of the expected cost in refining the estimate, and (2) at the business case gate, with a 25% more accurate estimate, reassess ROI. If the project no longer met the criteria, it was abandoned, and the feasibility investment was lost.

Table 30

Gated Agent (Beta Portfolio, F_{seed} = 50, No Risk Threshold, 25% Estimate Improvement)

		Investment in Phase 1 (Percentage of Expected Cost), I_{feas}						
		10%	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
	0%	-46	-46	-46	-46	-46	-46	-46
	10%	-46	-46	-46	-46	-46	-46	-46
	20%	-46	-46	-46	-46	-46	-46	-46
	40%	1622	**-28	1495	-34	-44	-41	-44
	60%	1442	1404	**-28	1316	-32	-35	-45
Expected ROI, ex-ante, ROI_{th}	**80%**	906	872	838	804	770	737	703
	100%	570	544	519	494	468	443	418
	120%	337	315	**-29	**-32	**-38	**-41	**-45
	140%	191	171	152	133	113	94	75
	160%	154	144	134	123	113	102	92
	180%	53	44	34	25	16	7	-2
	200%	++-5	++-11	++-16	++-22	++-27	++-32	++-38
	250%							

** early portfolio failure (path dependent loss) ++ ROI threshold too high

Notable in the results:

1. Higher returns were achievable despite the abandoned projects – 1622 returned for the Gated Agent vs 1191 for the Risk-Aware Agent.
2. Knowledge of the risk was not required.
3. A moderate investment of 10–30% allowed a lower ROI threshold (ROI_{th} > 40% to make a profit) compared to the ROI Agent (ROI_{th} > 80% to make a profit). The caveat was that the investment must significantly improve the cost estimate.

4. The highest returns (darkest green) tended to border the pink failure region, creating a drop-off ("cliff") between success and failure and a region of high sensitivity to selection criteria.

Two counterintuitive results require explanation. Some simulation runs (marked **) produced losses despite adjacent cells showing positive results. Code inspection and testing with varying seed funding levels revealed a path-dependent effect: actual costs were significantly higher than expected. Since feasibility investment was sunk regardless, the Gated Agent sometimes exhausted seed funding before making a profit. While untidy in presentation, this reflects a real phenomenon – early gambles can deplete capital, leaving only debt.

A second set of unexpected losses (marked ++) appeared along the bottom of Table 31, where the ROI threshold was high. Because the beta distributions were bounded, both cost estimates and actual costs fell within a known range. A high ROI threshold favoured low initial cost estimates, making post-feasibility estimates likely to rise. With higher cost estimates, projects were less likely to meet the ROI threshold. This occurred because the expected and actual costs followed a central tendency common in most probability distributions except uniform ones (Everitt, 2010) – see Table 31. In practice, portfolio managers selecting high-ROI projects are more likely to see cost estimates increase since low expected costs can inflate ROI projections.

Table 31

Gated Agent (Uniform Portfolio, F_{seed} = 50, No Risk Threshold, 25% Estimate Improvement)

		Investment in Phase 1 (Percentage of Expected Cost), I_{feas}						
		10%	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
Expected ROI, ex-ante, ROI_{th}	**0%**	-41	-42	-40	-41	-43	-46	-42
	10%	-41	-42	-40	-41	-43	-46	-42
	20%	-41	-42	-40	-41	-43	-46	-42
	40%	-3	-10	-27	-42	-16	-22	-28
	60%	-18	-18	-31	-34	-30	-43	-47
	80%	71	-32	-22	-26	-26	-32	-46
	100%	-31	-24	-21	-43	-40	-45	-44
	120%	-36	-20	-42	-27	-28	-33	-46
	140%	-31	-29	-41	-40	-44	-42	-27
	160%	64	-29	-42	-27	-41	-46	-25
	180%	55	-36	-36	-22	-44	-41	-45
	200%	-9	-41	-42	-41	-42	-46	-28
	250%	-24	-41	-41	-42	-41	-43	-42

Like the other agents, the first Gated Agent simulation performed poorly against the Uniform Portfolio. Only three portfolio runs made a profit, all at 10% feasibility investment. This indicates that a 25% improvement in cost accuracy is insufficient to guarantee profits with the Uniform Portfolio.

3.3.3.2 Gated Agent Results - 50% Improvement

The Gated Agent was rerun with a better cost estimate improvement (50%) at the end of feasibility. It was the only agent to show promise with the Uniform Portfolio (Table 32). The agent generated modest but positive portfolio returns, provided that a small initial investment (10–30%) yielded a substantially better cost estimate (in this case, a 50% improvement).

Table 32

Gated Agent (Uniform Portfolio, F_{seed} = 50, No Risk Threshold, 50% Estimate Improvement)

Expected ROI, ex-ante, ROI_{th}	Investment in Phase 1 (Percentage of Expected Cost), I_{feas}: 10%	20%	30%	40%	50%	60%	70%
0%	-41	-42	-40	-41	-43	-46	-42
10%	-41	-42	-40	-41	-43	-46	-42
20%	-41	-42	-40	-41	-43	-46	-42
40%	-36	-40	-40	-41	-41	-41	-41
60%	294	166	-42	-42	-40	-27	-44
80%	252	138	-20	-45	-25	-27	-44
100%	240	133	-43	-43	-40	-45	-44
120%	244	153	63	-27	-28	-46	-46
140%	124	45	-43	-42	-42	-31	-43
160%	-5	-42	-43	-30	-43	-41	-40
180%	-34	-36	-41	-43	-41	-43	-43
200%	-40	-41	-40	-43	-41	-43	-43
250%	-24	-41	-41	-42	-41	-43	-42

Notable from these results:

1. Selecting the correct thresholds was risky with the Uniform Portfolio, with only 11 of 91 approaches making a profit.
2. Under no conditions was the entire Uniform Portfolio passed over – the approach either succeeded or failed. This means the Gated Agent frequently “perceived wins” where none were to be had.

Similarly, for the Beta Portfolio (Table 33), improving the cost estimate at the business case gate to 50% accuracy improved the available options and maximum achievable portfolio value; that is, the region of green increased, and the returns increased. Notably, though, even with improved cost estimates, the general pattern of the results did not change significantly compared to the Gated Agent with only a 25% cost improvement. This is because there was still appreciable uncertainty in the revenue estimate, which was not improved. Perfecting the cost estimate only gets one so far.

Table 33

Gated Agent (Beta Portfolio, F_{seed} = 50, No Risk Threshold, 50% Estimate Improvement)

		Investment in Phase 1 (Percentage of Expected Cost), I_{feas}						
		10%	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**
	0%	-42	-43	-44	-45	-46	-47	-48
	10%	-42	-43	-44	-45	-46	-47	-48
	20%	1860	1758	1656	-44	-45	-43	-44
	40%	1717	-28	1537	-34	-44	-41	-44
	60%	1196	1124	-28	968	-32	-35	-45
Expected ROI, ex-ante, ROI_{th}	**80%**	791	738	685	-34	-43	-41	-44
	100%	423	384	344	305	266	227	-37
	120%	277	249	221	193	165	137	97
	140%	139	116	94	71	49	26	3
	160%	37	21	5	-41	-40	-41	-47
	180%	20	8	-3	-14	-25	-41	-47
	200%	-5	-11	-16	-22	-27	-32	-38
	250%							

3.3.4 Summary of Results

In the sections above, the individual results of each agent and portfolio combination were presented. Building on this, the performance of each agent is evaluated below and compared against other agents. Firstly, the success rate of each agent is compared. Then, the maximum portfolio value is compared along with the success criteria for each agent. Finally, the sensitivity of the parameters is considered.

3.3.4.1 Success Rates

The first performance measure was the success rate of each agent, which was defined as the number of profitable portfolio runs divided by the total number of portfolio runs. The success rate tells us how likely an agent was to generate a profit (any positive profit) given randomly chosen input parameters, that is, without knowledge of success criteria. For the Beta Portfolio, Table 34 tabulates the agent, simulation conditions, and the number of successes against the total number of portfolio trials (excluding portfolios where all projects were passed over) in the simulation grid for each agent. For the ROI agent, only eight trials met the condition $F_{seed} = 50$, of which half produced a profit. For the Risk Aware Agent, 50 trials resulted in 36 successes. For the Gated Agent, two scenarios were run, "Gated-25" with a 25% improvement in cost accuracy after feasibility (at the business case gate) and "Gated-50" with a 50% improvement in cost accuracy at the same gate. The success rates for the Gated Agent scenarios were similar: 21 of 36 (Gated-25) and 24 of 36 (Gated-50).

Table 34

Beta Portfolio Simulation Success Rate Summary (F_{seed} = 50, $ROI_{th} \geq 0$)

			Significance of Difference			
Agent	**Conditions**	**Success Rate**	**Gated-50**	**Gated-25**	**Risk A.**	**ROI**
ROI	Default	4 of 8 (50%)	0.625	0.333	0.788	0.000
Risk Aware	Default	36 of 50 (72%)	0.405	0.814	0.000	
Gated-25	25% cost accuracy improvement at gate 2	21 of 36 (58%)	0.535	0.000		
Gated-50	50% cost accuracy improvement at gate 2	24 of 36 (67%)	0.000			

Table 34 also shows (under the heading "Significance of Difference") the two-tailed binomial probability that the success rate of respective agents was unequal, that is, different. For example, the intersection of the ROI Agent (row) and Gated-50 Agent (column) indicates the probability of the success rates being unequal is 0.625 (62.5%). The significance of the difference in the success rates is useful because it clarifies whether a particular decision-making strategy was significantly better or worse than another.

The key observation from Table 34 for the Beta Portfolio is that since the significance of all the differences is below 90%. That is, none of the success rates differed at 90% confidence. Further, at 80% confidence, only the Gated-25 Agent and Risk Aware Agent differed significantly. This meant that, under the Beta Portfolio, a naive manager choosing randomly had approximately the same likelihood of making a profit regardless of the decision strategy (but no control over the size of that profit).

Table 35

Uniform Portfolio Simulation Success Rate Summary (F_{seed} = 50, $ROI_{th} \geq 0$)

			Significance of Difference			
Agent	**Conditions**	**Success Rate**	**Gated-50**	**Gated-25**	**Risk A.**	**ROI**
ROI	Default	0 of 9 (0%)	0.930	0.610	0.507	0.000
Risk Aware	Default	3 of 60 (5%)	0.999	0.417	0.000	
Gated-25	25% cost accuracy improvement at gate 2	3 of 39 (8%)	0.982	0.000		
Gated-50	50% cost accuracy improvement at gate 2	11 of 39 (28%)	0.000			

With the Uniform Portfolio (Table 35), the ROI agent had no successes, and the Risk Aware agent fared only a little better, with three successes out of 60 trials. The Gated-25 Agent also had three successes but attempted fewer trials, and the Gated-50 Agent performed best with 11 successes out of 39 trials (28%). However, a major difference compared to the Beta Portfolio is the significance of the difference between the Gated-50 Agent and all the other agents (> 90%). This demonstrates that, under high uncertainty, if a feasibility study can significantly improve cost estimates alone, it becomes possible to profit. This further strengthens the argument supporting decision gates and the objective application of decision criteria.

3.3.4.2 Maximum Profit and Success Criteria

The second performance measure was maximum profit, which is summarised and tabulated here, beginning with the criteria (regions) in which the agents were successful in Table 36, and observations are noted in the following text.

With the Beta Portfolio, all agents could make a profit. However, each agent had a different approach to risk. With the ROI Agent, selecting a high ROI threshold ($ROI_{th} \geq 80\%$) was the only way to proxy for risk. In contrast, the Risk-Aware Agent was able to address risk more directly by setting a threshold for the coefficient of variation ($CV_{th} \leq 0.36$) and consequently selecting a more reasonable ROI threshold ($ROI_{th} \geq 40\%$). This gave access to more projects and generated a slightly higher return. However, the real gains were made by the Gated-25 and Gated-50 agents, which ignored risk ex-ante and instead engaged with the project through a feasibility study, which yielded improved cost estimates. Provided that the feasibility investment was relatively low ($I_{feas} \leq 30\%$), a significantly higher profit (than the ROI and Risk-Aware agents) was achievable.

Table 36

*Portfolio Success Criteria**

	ROI Agent	Risk-Aware Agent	Gated-25 Agent	Gated-50 Agent
Beta Portfolio	$ROI_{th} \geq 80\%$	$ROI_{th} \geq 40\%$ $CV_{th} \leq 0.36$	$ROI_{th} \in [40, 160]$ % $I_{feas} \leq 30\%$	$ROI_{th} \in [20, 140]$ % $I_{feas} \leq 30\%$
Max Profit	1082	1191	1622	1860
Uniform Portfolio	No identifiable success region	No identifiable success region	No identifiable success region	$ROI_{th} \in [60, 140]$ % $I_{feas} \leq 30\%$
Max Profit	-9	124	71	294

* approximate regions most likely to deliver successful results, see individual results above. All with F_{seed} = 50.
ROI_{th} = ROI threshold CV_{th} = risk threshold I_{feas} = feasibility investment F_{seed} = seed funding

Conversely, with the Uniform Portfolio, only the Gated-50 Agent could clearly define success criteria for operating in this environment and make a relatively modest profit. This came on the condition that the cost estimate was 50% closer to the real cost at the end of the feasibility study.

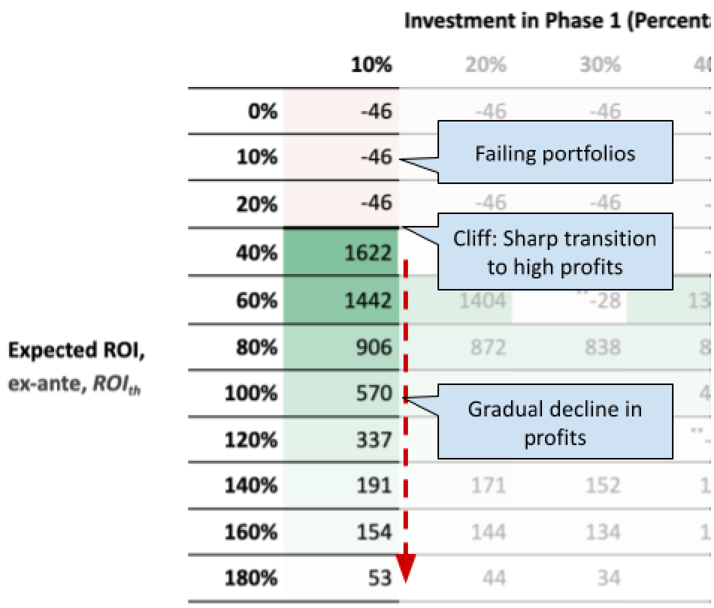
For the Gated agents, it is notable that the success criteria always required the feasibility investment to be less than or equal to 30% ($I_{feas} \leq 30\%$). This avoided overspending on the feasibility study and ensured the bulk of the budget allocated to a project could be reallocated if it was abandoned. Generally, the less spent on a feasibility study (to yield the same cost estimate improvement), the better, as was evident in the Gated Agent's results.

3.3.4.3 Parametric Sensitivity

A pattern of "cliffs" was observed in the results and is illustrated in Figure 25 below. Moving from top to bottom, initially, with low ROI thresholds, the portfolios failed because seed funding was exhausted. However, as the ROI threshold increased, a sharp transition from portfolio failure to high (or even maximum) profits was observed, followed by a slow decline in profits as the ROI threshold increased. This occurs because the probability of completing the initial projects increases with increasing ROI threshold. At low ROI thresholds, there were too many "bad projects", so the seed funding was exhausted. Once the small seed funding allocation (F_{seed} = 50) had grown, a new condition prevailed: The lower the ROI threshold, the more projects were selected, and thus, greater profits were possible. As the ROI threshold increased, the additional selectivity gradually reduced overall portfolio profits. From a naive perspective, the lower the ROI threshold, the better, but this assumes knowledge of the location of the cliff, so it would be better guidance to select a slightly higher ROI threshold to include a margin for error and avoid early portfolio failure.

Figure 25

Portfolio "Cliffs"



3.4 Discussion

The simulation results above show that the Gated Agent performed better than the ROI and Risk-Aware Agents when executing the Uniform Portfolio and had higher earning potential when executing the Beta Portfolio. From these results, it is reasonable to conclude that R&D processes using decision gates will perform better under higher uncertainty or where the underlying probability distribution is not amenable to prediction, provided that managers are willing to make objective go/no-go decisions. It is the nature of the uniform distribution to have no defined mode – it can be any value in the distribution's range – yet the Gated Agent produced statistically significant results against the Uniform Portfolio. This illustrates the value of feasibility studies and highlights the importance of stage-gates in new product development. Snowden (2005) suggests that complex and chaotic systems lack predictability and must be engaged to make any learnings or gains, as was demonstrated with the Uniform Portfolio, where analytical approaches (ROI and risk) failed, and even the experimental approach (using stage-gates) struggled to make ground in most scenarios.

In R&D, there is no right tool for the job because the tool (ROI, risk, stage-gates) and environment (business, technology, market) will affect the nature of the R&D system and the level of uncertainty. Loch (2000) concludes, "There is no 'best practice' NPD process. Rather, a company should develop a customised R&D project portfolio and a corresponding mixture of processes, which together meet its strategic innovation needs." For example, in SaaS web development, the technological complexity is low due to well-established tools and platforms, but the nature of target markets can vary immensely. Not that it is an either-or consideration – all the simulated approaches can be used together. However, the demonstration here shows that using strict go/no-go gates to implement and execute based on predetermined criteria (the only tested approach that engages with and explores the system) was the most effective overall because it requires the least information about the environment ex-ante. Committing up-front required that cost and revenue were somewhat predictable, whereas the business case gate allowed for an engage-and-learn approach.

Box 1

The Feasibility Rule

For a feasibility phase to add value, it should cost less than 30% of the project budget and significantly improve cost estimates.

The robustness of the stage-gate approach was evident in its ability to generate returns in unpredictable environments (the Uniform Portfolio) and its relative insensitivity to the accuracy of cost estimates. The approach implemented here assumed that valuable information would become available during the feasibility phase. That's the implicit gamble in a feasibility phase: Managers bet 10–30 % of the estimated project cost to reduce the error in the cost estimate, often from [-50, +80] % to as low as ± 15 %. So managers replace one gamble with another: They no longer bet on the whole project; instead, they bet that they can learn about (de-risk) the project for a fraction of the expected cost. For the same seed funding, this allows many more small bets (feasibility studies) to be run, which is why gates are helpful for SMEs on restricted budgets. This gave rise to the Feasibility Rule in Box 1, highlighting that feasibility is not just about identifying what problems are technically solvable; it is about resolving uncertainty into manageable risks so estimates are improved and the project viability can be better assessed.

In the paragraph above, "They no longer bet on the whole project" implies a seemingly unreasonable degree of naivety. It is doubtful that anyone would intentionally bet on the *whole project* without some preliminary investigation. But, remember that the design of the ROI and Risk-Aware agents was to implement commitment escalation by going "all in" from the start, so they model "pure dogged commitment" to project completion. This might seem unrealistic, but it is common in literature, with all 22 included articles in the brief review (Table 22, page 84) considering only project selection without the possibility of abandoning it. Further, the complex set of beliefs and biases can easily lead managers to early overcommitment, thus, while the modelling choices for the ROI and

Risk-Aware agents represent idealised extremes, they also capture realistic tendencies observed in practice. Conversely, the Gated Agent simulations implement objectivity at the business case gate and are unconcerned with shame, embarrassment, or impact on reputation. If the no-go criteria were met, the project was abandoned – end of story. The behaviour of most managers is likely to fall on a continuum between dogged commitment and pure objectivity, so the agents usefully bracket human behaviour, as do the results.

With real people, there is a bias toward escalation because most of us are compelled to demonstrate integrity, to "say what we do, and do what we say." (Cialdini, 1993). This dependability is useful in establishing and maintaining trust but also contains a trap for the uninitiated. When managers say, with expert confidence, "we can achieve X", they leave no room to back out. Conversely, if they take the stage-gate approach and negotiate no-go criteria up front, they can say, "We will explore X, and if X is not viable, we will abandon it." This is why up-front contracting about the process and criteria is so important – so that managers do what they said they would do when they abandon a project: If no-go criteria and actions are negotiated up-front, then managers effectively reserve the right to enact them in a consistent, congruent and trustworthy way.

The tendency for commitment escalation highlights a key risk in managing complex systems – pushing processes to their limits can lead to systemic failure (Snowden, 2005). This manifested as a pattern in the simulation results where a sharp transition between good profits and failing portfolios (cliffs) was observed. There is a warning here for managers: Flying too close to the sun could result in disaster. Snowden (2005) observes that complex systems may become chaotic when managed using best practice (read: inflexible) approaches. Similarly, attempting to maximise the return in R&D by repeatedly tuning and systematising the development model might push the operating point against the precipice and risk failure and losses. Practitioners are advised to maintain a degree of agility and a margin for error.

The simulations deliberately omitted several factors. These omissions create opportunities to improve the models in the future. Firstly, evaluation of projects was limited to yes/no in a specific sequence. Implementing scoring models, selection matrices, and force ranking of projects would provide insight into the viability of different selection approaches. However, the absence of these project selection approaches did not limit the utility of the simulation when comparing agents against each other – the design was that each agent had to perform in the *same* environment.

A second omission was that the project portfolios were static in that they did not include any dynamic element to add more realistic complexity from feedback or change. Instead, uncertainty was modelled through the underlying probability distributions. While it was not possible for the agents to know the actual cost and revenue up front, this is subtly different from dealing with change as the project proceeds. However, from the agents' perspectives, there was no difference in whether the uncertainty came from an unknown or change, and the former was simpler to implement on a comparative basis. Further, additional complexity from feedback or change would have eroded the comparative nature of the simulation and added uncertainty to the results, making them more difficult to interpret.

A third omission was that resources were assumed to be available as needed (except seed funding). Capacity and capability were not implemented in the simulation, yet most simulation frameworks incorporate the concept of resource constraints and queuing. Adding these would add a dimension of complexity to the model but might also shed light on the project selection process. Initial project selection was not a focus of this paper, but future work in this space might also be beneficial.

The emergence of the Feasibility Rule and the validation of the use of objective decision gates in the Framework fulfils the aims of Part B. The first aim was to validate the Framework by illustrating its application compared to other decision-making approaches. This was achieved by demonstrating that objectively facilitated decision gates improved outcomes in the simulated R&D portfolio, particularly under high levels of uncertainty. Many factors presented in the Framework were included in the simulations: A clear strategy (maximise portfolio value) was defined, and capabilities such as agility (willingness to abandon) and project tracking were implicit in the simulation accounting. Precise go/no-go criteria were pre-committed and executed with rigour at the decision gates, and budgets were allocated in the form of portfolio seed funding. Gatekeeper independence was implicit in the form of computational objectivity, and costs, perceived benefits and market value were encoded in the probability distributions representing the projects. Also represented in the probability distributions were risks and uncertainty relating to assumption validity, technical feasibility, data quality issues from inadequate data and estimation error, and change from market volatility. The comprehensive coverage of the Framework's factors increases confidence that the simulation results provide a legitimate means of validating the Framework. Finally, by validating the Framework, the simulation outcomes support the extant literature regarding using decision gates in R&D.

The second aim was to test the sensitivity of the outcomes to varying decision-making strategies and feasibility phase investment levels. The Feasibility Rule directly addresses this aim, suggesting that the investment in the feasibility phase should be kept below 30% of the expected project cost. Further, the simulations showed that using decision gates was more likely to provide benefits under higher levels of uncertainty, where a strategy of engaging and learning was more suitable than up-front design and commitment.

3.5 Conclusion of Simulation

In Part B, a simulation model was developed to test the Framework in semi-realistic product development scenarios (the Beta Portfolio) and under high uncertainty (the Uniform Portfolio). This novel approach incorporated mid-project decision gates alongside up-front project selection criteria – the latter being the prevailing focus in the literature. The purpose of the simulation was to illustrate the validity of the Go/No-Go Framework, which is designed as "a formal process that ensures managers make strategically aligned decisions by evaluating options objectively using distancing strategies and multiple lenses." To that end, the simulation encoded formality, objectivity, clear gate criteria, and the ability to evaluate options through feasibility studies.

The findings validate the Framework and contribute empirical justification for the use of decision gates in R&D. Using clearly defined gate criteria allows for early risk management and identification of go/no-go options. Additional contributions from the simulation study include: (1) guidance on where decision gates perform best (in high-uncertainty environments), (2) practical insights into the design of feasibility studies – specifically that they should consume 10–30% of total effort and deliver refined cost estimates, (3) the identification of cliffs – sharp boundaries between highly successful and failing portfolios – which help practitioners avoid overly ambitious criteria selection, and (4) the novel simulation of mid-project decision gates, a less explored area in existing research.

The nature of the simulation, while useful for comparison, was unrealistic in some ways, and these constitute areas for further research. Firstly, the project selection approach was limited to a simplistic "yes – go ahead" or "no – next project" decision. Second, projects were executed sequentially in the same order and never in parallel. Third, resource constraints were ignored. These simplifications allowed the simulation to assume that resources were always available, which is often untrue in reality. Future research could implement a resource constraint and run as many projects as possible in paralle to extend the goal from maximising profit to maximising profit in the shortest possible time. This would also add feedback and complexity between the projects, making prediction more difficult.

A final area of additional research came from identifying the strong bias in the simulation literature toward initial project selection rather than mid-project evaluation. This affirmed the originality of this thesis and highlights the need for additional quantitative analysis of what happens as projects proceed and when to abandon them – not just how to select them in the beginning.

4 Integrated Conclusion – Literature and Simulation

This study sought to address the question:

RQ: How should firms evaluate whether to continue or abandon projects in high-tech R&D?

Part A provided an integrative literature review using a novel AI-assisted search method. The aim was to identify those key factors involved in go/no-go decisions from which to develop a conceptual model that is presented as a decision-making framework. While focused intensely on go/no-go decisions, the review spanned the breadth of the R&D literature. The Go/No-Go Framework offers practical guidance while providing academics with a thorough AI-assisted review of the topic.

The Framework elucidates the details of a theoretical go/no-go decision gate, by capturing the constructs (approach, antecedents, tools, criteria and sources of error) and factors relevant to good go/no-go decision-making. Central to this is the recommended approach synthesised into the Framework from the literature, which is: "A formal, staged process, focused on early learning, with clear gate criteria, that identifies and retains strategic options." Judicious application of the Framework and the ability to avoid commitment escalation and abandon a project early is essential for SMEs who do not have access to large quantities of seed funding because this allows limited budgets to be used to explore more potential projects. Although the simulation results suggest larger firms might be less susceptible to portfolio failure, they are also expected to benefit from using the Framework to guide R&D decision-making. A structured approach to decision-making should benefit all firms.

Part B built on the Go/No-Go Framework, which was based on the literature review, to develop a novel simulation model that illustrated the value of decision gates compared to up-front risk and ROI-based approaches, particularly in complex and chaotic environments. This was done to test the key concepts of the Framework and illustrate its application. Three decision-making strategies were compared – ROI-based, risk-based and gated (reassessed after a feasibility stage), and it was found that the strategy of investing in a feasibility study followed by a hard-nosed go/no-go decision gate outperformed the ROI-based and risk-based strategies. The simulation outcomes supported the use of decision gates generally and, by induction, the proposed Go/No-Go Framework. Decision gates were more beneficial in environments where risk was difficult to assess, and the ROI-based and risk-based strategies failed under such conditions.

One of the most actionable outcomes was the Feasibility Rule: *For a feasibility phase to add value, it should cost less than 30% of the project budget and significantly improve cost estimates.* This rule gives managers a practical heuristic for deciding how much to invest in a project's feasibility stage.

The main theoretical contribution of this research was the literature review and Framework, which added to the body of knowledge on R&D decision-making by focusing specifically on go/no-go decisions. The literature review integrated extant theories of R&D decision-making by reviewing relevant research on real options, strategic management, decision-making, commitment escalation and biases, portfolio management, and project management. This resulted in a Framework that captured the essential constructs and factors of go/no-go decision-making without overwhelming the reader with too much information or detail.

For practitioners, the Framework provides a tool for evaluating process maturity in portfolio management and project governance and a checklist for sound decision-making. It has the potential to inform R&D portfolio governance practices and support policymakers in determining project funding criteria and guidelines and go/no-go conditions. Practitioners are encouraged to watch for biases that lead to commitment escalation and seek independent perspectives in their portfolio decision-making. No universal theory of go/no-go decision-making was discovered because the application of the Framework will be highly situational. However, this work did identify some clear guidelines for practitioners:

Table 37

The Do's and Don'ts of Go/No-Go Decision Making

Do	Don't
Have a clear strategy	Chase any/all opportunities
Contract options and go/no-go criteria up-front	Wait until it is going badly to negotiate
Use a formal decision-making process	Overanalyse, or just go with your gut
Use independent gatekeepers	Let invested parties make portfolio decisions

An additional and significant contribution was the exemplification of the use of AI to facilitate literature reviews. Using AI-determined relevance, a novel search method was developed to rate thousands of articles. This approach avoided manual filtering and enabled iteration on the research question by making searching very fast. Although requiring human supervision and further validation, the widespread application of this approach has the potential to accelerate research on a global scale and enable large-scale integration of existing research. This work adds credibility to the use of AI for this purpose.

This study has three main limitations. The first and most fundamental limitation is the framework's lack of real-world empirical validation. The simulation approach was presented as illustrative (rather than an empirical study) because it simplified decision-making by considering only rational factors (cost, risk, return) and omitted the human factors (commitment escalation, biases and heuristics) identified in the Framework. Second, while evident in the simulation results, the Feasibility Rule might also not translate into the real world. Third, the AI-assisted method has yet to be proven or generally accepted. Because of this, relevant articles may have been omitted – this is addressed empirically in Appendix G.

This study has identified numerous directions for future research as set out below:

Quantify the failure rate in high-tech SMEs: While there is some reporting on R&D project failure rates, the statistics vary widely by industry. It is difficult to ascertain what the failure rate in high-tech SMEs is from the available data. Future empirical research can work to clarify this and support policymakers to take appropriate corrective action.

Real-world testing of the Go/No-Go Framework: Beyond simulation, the next logical step is empirically validating the Framework. Starting small might imply case studies or grounded research to explore and iterate the Framework and its application in real-world scenarios. At a larger scale, the Framework could be tested empirically by deploying it in a sample of organisations while maintaining a control sample and performing statistical testing to determine its efficacy. Alternatively, SMEs might have historical data to use as a baseline against which to measure improvements in R&D investments.

Develop an assessment tool (a health check) based on the Framework and test it empirically to validate the Framework and its sensitivity to decision-making factors. This tool would add value by providing academics and practitioners with a way to measure the effectiveness of an R&D decision-making process. In SMEs, this also has implications for governance structure.

Validate the Feasibility Rule with further research by using surveys, empirical studies, or further simulation and quantitative analysis to elaborate on the rule's applicability and add substance to right-sizing the feasibility phase of R&D projects.

Simulation extensions: There are some incremental opportunities to improve the simulation. One is to extend it to include a decision-making agent that considers more than one project at a time (perhaps three) and selects the best project. This would implement a more realistic project selection scenario against which decision-making strategies could be tested. Another is to apply capacity constraints to the model to determine if limiting the human resources on R&D projects has implications for the most appropriate decision-making strategy. A third is to incorporate the impact of bias into the simulation, bringing the purely programmatic approach closer to incorporating behavioural economic theory and highlighting which decision-making approaches are more or less sensitive to human bias.

AI-assisted decision-making: A more significant leap is extending the simulation to add an “AI Agent” to explore the efficacy of AI or machine learning to support decision-making. AI-assisted decision-making is an emerging field (Steyvers & Kumar, 2023). It can add value by augmenting human decision-making with a more rational perspective, potentially even implementing gatekeeper independence in the form of AI. This might be augmented by empirical studies that test the efficacy of (1) human, (2) AI and (3) AI-assisted human decision-making.

Integrate emerging fields: Beyond AI-assisted decision-making, other relevant fields of study have advanced beyond the R&D literature. Decision-making (Kahneman, 2011) and group intelligence (Woolley et al., 2015) are of particular note. These key fields of study directly affect decision-making and might reveal relevant factors and improvements to the Framework developed here and for new product development.

Validate the AI-assisted search method more thoroughly using multi-rater-based empirical methods. The method was applied to one topic in management science here. However, it is unknown how it will perform in domains such as mathematics and physics, where concepts are more specialised and expressed differently. The method must be validated in each domain and against many different research questions to be proven.

Through the review and simulation, this thesis has advanced the field of decision-making in new product development and exemplified the use of AI in literature reviews. By developing a novel framework that focuses specifically on go/no-go decisions, this work emphasises the importance of good R&D project and portfolio management. Practitioners are urged to learn and apply the concepts embodied in the Framework, and academics are urged to explore the use of AI in their research. While limitations and opportunities for future research remain, this work provides practitioners with the tools to implement effective decision gates, empowering project, product and portfolio managers to utilise precious resources to maximise the benefits to people, firms and society.

5 References

Abd-Alrazaq, A., Schneider, J., Mifsud, B., Alam, T., Househ, M., Hamdi, M., & Shah, Z. (2021). A comprehensive overview of the COVID-19 literature: Machine learning–based bibliometric analysis. *Journal of Medical Internet Research, 23*(3), e23703.

Abyad, A. (2020). The Pareto principle: Applying the 80/20 rule to your business. *Middle East Journal of Business, 15*(1), 6–9. https://doi.org/10.5742/MEJB.2021.93880

Acebes, F., Poza, D., González-Varona, J. M., & López-Paredes, A. (2022). Stochastic earned duration analysis for project schedule management. *Engineering, 9*, 148–161.

Aguilar-Rivera, A., & Valenzuela-Rendón, M. (2017). A new multi-period investment strategies method based on evolutionary algorithms. *Neural Computing and Applications, 31*(3), 923–937. https://doi.org/10.1007/s00521-017-3121-6

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more* (1st ed.). Hyperion.

AppSheet. (n.d.). *AppSheet*. Retrieved March 1, 2025, from https://about.appsheet.com/home/

Australian Business Deans Council. (2023). *ABDC journal quality list*. Australian Business Deans Council. Retrieved July 16, 2023, from https://abdc.edu.au/abdc-journal-quality-list/

Ayala-Cruz, J. (2016). Project risk planning in high-tech new product development. *Academia Revista Latinoamericana de Administración, 29*(2), 110–124. https://doi.org/10.1108/arla-11-2015-0297

Bansal, P., Bertels, S., Ewart, T., MacConnachie, P., & O'Brien, J. (2012). Bridging the research–practice gap. *Academy of Management Perspectives, 26*(1), 73–92. https://doi.org/10.5465/amp.2011.0140

Bard, J. F., & Feinberg, A. (1989). A two-phase methodology for technology selection and system design. *IEEE Transactions on Engineering Management, 36*(1), 28–36. https://doi.org/10.1109/17.19980

Barucke Marcondes, G. A., Leme, R. C., Leme, M. da S., & Sanches da Silva, C. E. (2016). Using mean-Gini and stochastic dominance to choose project portfolios with parameter uncertainty. *The Engineering Economist, 62*(1), 33–53. https://doi.org/10.1080/0013791x.2016.1176283

Behrens, J., & Ernst, H. (2013). What keeps managers away from a losing course of action? Go/stop decisions in new product development. *Journal of Product Innovation Management, 31*(2), 361–374. https://doi.org/10.1111/jpim.12100

Bolaños, F., Salatino, A., Osborne, F., & Motta, E. (2024). Artificial intelligence for literature reviews: Opportunities and challenges. *Artificial Intelligence Review, 57*(10), 259. https://doi.org/10.1007/s10462-024-10902-3

Boulding, W., Morgan, R., & Staelin, R. (1997). Pulling the plug to stop the new product drain. *Journal of Marketing Research, 34*(1), 164–176. https://doi.org/10.1177/002224379703400114

Brandimarte, P. (2014). *Handbook in Monte Carlo simulation: Applications in financial engineering, risk management, and economics*. John Wiley & Sons.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review, 17*(1), 39–61. https://doi.org/10.5465/amr.1992.4279568

Buehler, R., Griffin, D., & Ross, M. (2002). Inside the planning fallacy: The causes and consequences of optimistic time predictions. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 250–270). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.016

Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. (2023). On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management, 26*(7), 233–241.

Castellion, G., & Markham, S. K. (2013). Myths about new product failure rates: Influence of *argumentum ad populum* and self-interest. *Journal of Product Innovation Management, 30*(5), 976–979. https://doi.org/10.1111/j.1540-5885.2012.01009.x

Chan, T., Nickerson, J. A., & Owan, H. (2007). Strategic management of R&D pipelines with cospecialized investments and technology markets. *Management Science, 53*(4), 667–682. https://doi.org/10.1287/mnsc.1060.0676

Cheung, M. T., Greenfield, P. F., & Liao, Z. (2009). Selecting R&D projects for technology-based innovation. *Journal of General Management, 35*(2), 61–80. https://doi.org/10.1177/030630700903500205

Chiu, C. C., & Lin, K. S. (2019, July). Rule-based BCG matrix for product portfolio analysis. In *International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing* (pp. 17–32). Springer International Publishing.

Choungsirakulwit, J., & Sutivong, D. (2007). Portfolio management of option-based investment in technology research and development. In *6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)* (pp. 732–737). IEEE. https://doi.org/10.1109/icis.2007.150

Cialdini, R. B. (1993). *Influence: The psychology of persuasion* (Rev. ed.). Morrow.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Coldrick, S., Longhurst, P., Ivey, P., & Hannis, J. (2005). An R&D options selection model for investment decisions. *Technovation, 25*(3), 185–193. https://doi.org/10.1016/S0166-4972(03)00099-3

Conversano, C., & Lizzeri, A. (2012). Conditional performance attribution for equity portfolio. In *Mathematical and Statistical Methods for Actuarial Sciences and Finance* (pp. 105–113). https://doi.org/10.1007/978-88-470-2342-0_13

Cooper, R. G. (1990). Stage-gate systems: A new tool for managing new products. *Business Horizons, 33*(3), 44–54.

Cooper, R. G. (1999). From experience: The invisible success factors in product innovation. *Journal of Product Innovation Management, 16*(2), 115–133.

Cooper, R. G. (2006). Managing technology development projects. *Research-Technology Management, 49*(6), 23–31. https://doi.org/10.1080/08956308.2006.11657405

Cooper, R. G. (2019). The drivers of success in new-product development. *Industrial Marketing Management, 76*, 36–47. https://doi.org/10.1016/j.indmarman.2018.07.005

Cooper, R. G. (2022). The 5th generation stage-gate idea-to-launch process. *IEEE Engineering Management Review, 50*(4), 43–55. https://doi.org/10.1109/EMR.2022.3222937

Cooper, R. G., & Sommer, A. F. (2023). Dynamic portfolio management for new product development. *Research-Technology Management, 66*(3), 19–31.

Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (1997). Portfolio management in new product development: Lessons from the leaders – I. *Research-Technology Management, 40*(5), 16–28.

Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (2000). New problems, new solutions: Making portfolio management more effective. *Research-Technology Management, 43*(2), 18–33. https://doi.org/10.1080/08956308.2000.11671338

Crawford, C. M. (1987). New product failure rates: A reprise. *Research Management, 30*(4), 20–24. https://doi.org/10.1080/00345334.1987.11757047

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cronin, M. A., & George, E. (2023). The why and how of the integrative review. *Organisational Research Methods, 26*(1), 168–192. https://doi.org/10.1177/1094428120935507

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Prentice Hall.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. *Academy of Management Review, 32*(2), 480–499.

Del Campo, C., Pauser, S., Steiner, E., & Vetschera, R. (2016). Decision-making styles and the use of heuristics in decision making. *Journal of Business Economics, 86*, 389–412.

Dhillon, R., & Nguyen, Q. C. (2021). *Strategies to respond to a VUCA world*.

Dorison, C. A., Umphres, C. K., & Lerner, J. S. (2022). Staying the course: Decision-makers who escalate commitment are trusted and trustworthy. *Journal of Experimental Psychology: General, 151*(4), 960. https://doi.org/10.1037/xge0001101

Dufault, S., Wason, J. M. S., & Mozgunov, P. (2023). A flexible multi-metric Bayesian framework for decision-making in Phase II multi-arm multi-stage studies. *arXiv preprint arXiv:2302.07290.* https://arxiv.org/abs/2302.07290

Eliens, R., Eling, K., Gelper, S., & Langerak, F. (2018). Rational versus intuitive gatekeeping: Escalation of commitment in the front end of NPD. *Journal of Product Innovation Management, 35*(6), 890–907. https://doi.org/10.1111/jpim.12452

Elsevier. (2024). *Scopus AI: Trusted content. Powered by responsible AI.* Retrieved October 5, 2024, from https://www.elsevier.com/products/scopus/scopus-ai

Everitt, B. S. (2010). *The Cambridge dictionary of statistics* (4th ed.). Cambridge University Press.

Farshchian, M. M., & Heravi, G. (2018). Probabilistic assessment of cost, time, and revenue in a portfolio of projects using stochastic agent-based simulation. *Journal of Construction Engineering and Management, 144*(5). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001476

Ferrati, F., Kim, P. H., & Muffatto, M. (2024). Generative AI in entrepreneurship research: Principles and practical guidance for intelligence augmentation. *Foundations and Trends® in Entrepreneurship, 20*(3), 245–383.

Ford, S., Aubert, C., & Ryckewaert, E. (2016). *Reducing the risk of failure in new product development: Getting it right at the front end of innovation*. Institute for Manufacturing, University of Cambridge. https://www.ifm.eng.cam.ac.uk/uploads/Resources/Reports/Reducing-the-risk-of-failure-in-new-product-development-report.pdf

Galasso, A., & Simcoe, T. S. (2011). CEO overconfidence and innovation. *Management Science, 57*(8), 1469–1484.

Garg, S. (2020). Venture governance: A new horizon for corporate governance. *Academy of Management Perspectives, 34*(2), 252–265. https://doi.org/10.5465/amp.2017.0178

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology, 75*(6), 728.

Gilbert-Saad, A., Siedlok, F., & McNaughton, R. B. (2018). Decision and design heuristics in the context of entrepreneurial uncertainties. *Journal of Business Venturing Insights, 9*, 75–80. https://doi.org/10.1016/j.jbvi.2018.02.006

GitHub. (n.d.). *GitHub Codespaces*. GitHub. https://github.com/features/codespaces

Google. (2024). *Home - Apps Script*. Google for Developers. https://developers.google.com/apps-script

Gupta, P., Nguyen, T. N., Gonzalez, C., & Woolley, A. W. (2023). Fostering collective intelligence in human–AI collaboration: Laying the groundwork for COHUMAIN. *Topics in Cognitive Science.* https://doi.org/[insert DOI if available]

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (4th ed.). Advanced Analytics, LLC.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon Books.

Harvey, N. (2020). Use of heuristics: Insights from forecasting research. In *Judgement and choice: Perspectives on the work of Daniel Kahneman* (pp. 5–24). Psychology Press. https://doi.org/10.1080/13546780600872502

Hasan, R. U., & Chyi, T. M. (2017). Practical application of balanced scorecard – A literature review. *Journal of Strategy and Performance Management, 5*(3), 87.

Hauser, J. R., & Zettelmeyer, F. (1997). Metrics to evaluate R, D & E. *Research-Technology Management, 40*(4), 32–38. https://doi.org/10.1080/08956308.1997.11671140

Havlík, V. (2023). Meaning and understanding in large language models. *arXiv preprint arXiv:2310.17407.* https://arxiv.org/abs/2310.17407

Henderson, B. D. (1970). The product portfolio. *The Boston Consulting Group Perspective, 66.*

Huchzermeier, A., & Loch, C. H. (2001). Project management under risk: Using the real options approach to evaluate flexibility in R&D. *Management Science, 47*(1), 85–101. https://doi.org/10.1287/mnsc.47.1.85.10661

Hutchison-Krupat, J., & Kavadias, S. (2015). Strategic resource allocation: Top-down, bottom-up, and the value of strategic buckets. *Management Science, 61*(2), 391–412.

Ilie-Zudor, E., & Monostori, L. (2009). Agent-based framework for pre-contractual evaluation of participants in project-delivery supply chains. *Assembly Automation, 29*(2), 137–153. https://doi.org/10.1108/01445150910945598

Jagle, A. J. (1999). Shareholder value, real options, and innovation in technology-intensive companies. *R&D Management, 29*(3), 271–288. https://doi.org/10.1111/1467-9310.00136

Jin, Q., Leaman, R., & Lu, Z. (2024). PubMed and beyond: Biomedical literature search in the age of artificial intelligence. *EBioMedicine, 100*, 104988. https://doi.org/10.1016/j.ebiom.2024.104988

Joos, L., Keim, D. A., & Fischer, M. T. (2024). Cutting through the clutter: The potential of LLMs for efficient filtration in systematic literature reviews. *arXiv preprint arXiv:2407.10652.*

Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The use of artificial intelligence in writing scientific review articles. *Current Osteoporosis Reports, 22*(1), 115–121.

Kahneman, D. (2011). *Thinking, fast and slow.* Allen Lane.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430–454.

Kahneman, D., & Tversky, A. (1977). *Intuitive prediction: Biases and corrective procedures* (p. 0044). Decision Research.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. Little, Brown Spark.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review, 70*(1), 71–79.

Kauffman, R. J., Liu, J., & Ma, D. (2015). Technology investment decision-making under uncertainty. *Information Technology and Management, 16*, 153–172. https://doi.org/10.1007/s10799-014-0212-2

Kazak, E., & Pohlmeier, W. (2019). Testing out-of-sample portfolio performance. *International Journal of Forecasting, 35*(2), 540–554. https://doi.org/10.1016/j.ijforecast.2018.09.010

Kettunen, J., & Salo, A. (2017). Estimation of downside risks in project portfolio selection. *Production and Operations Management, 26*(10), 1839–1853. https://doi.org/10.1111/poms.12727

Kholany, M. M. S. E., & Abdelsalam, H. M. (2017). Multi-objective binary cuckoo search for constrained project portfolio selection under uncertainty. *European Journal of Industrial Engineering, 11*(6), 818–841. https://doi.org/10.1504/EJIE.2017.089107

Killen, C. P., Hunt, R. A., & Kleinschmidt, E. J. (2008). Project portfolio management for product innovation. *International Journal of Quality & Reliability Management, 25*(1), 24–38. https://doi.org/[insert DOI if available]

Kim, Y.-H., Park, S.-W., & Sawng, Y.-W. (2016). Improving new product development (NPD) process by analysing failure cases. *Asia Pacific Journal of Innovation and Entrepreneurship, 10*(1), 134–150. https://doi.org/10.1108/apjie-12-2016-002

Knight, F. H. (1921). *Risk, uncertainty and profit* (Vol. 31). Houghton Mifflin.

Kohn, S., Steinkellner, T., & Thome, T. (2021). Dealing with market uncertainty in product development: A systematic literature review. *PMC.* Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC10788447/

Koller, T., Goedhart, M., & Wessels, D. (2010). *Valuation: Measuring and managing the value of companies*. John Wiley & Sons.

Landschaft, A., Antweiler, D., Mackay, S., Kugler, S., Rüping, S., Wrobel, S., ... & Allende-Cid, H. (2024). Implementation and evaluation of an additional GPT-4-based reviewer in PRISMA-based medical systematic literature reviews. *International Journal of Medical Informatics, 189*, 105531.

Lee, J. M., Park, J. C., & Chen, G. (2022). A cognitive perspective on real options investment: CEO overconfidence. *Strategic Management Journal, 44*(4), 1084–1110. https://doi.org/10.1002/smj.3469

Lee, P. Y., Salim, H., Abdullah, A., & Teo, C. H. (2023). Use of ChatGPT in medical research and scientific writing. *Malaysian Family Physician, 18*, 58. CLOCKSS. https://doi.org/10.51866/cm0006

Legkokonets, V. A., Vasilenko, A. A., & Khasanov, M. M. (2020, October). Selection of optimal developing strategy for a group of prospects in terms of geological uncertainties and infrastructure constraints. In *SPE Russian Petroleum Technology Conference* (p. D043S022R001). Society of Petroleum Engineers.

Levitt, R. E. (2011). Towards project management 2.0. *Engineering Project Organization Journal, 1*(3), 197–210. https://doi.org/10.1080/21573727.2011.609558

Li, Y., & Zhang, S. (2020). Modeling and simulation of firm-leading self-organization evolution process in open source product design project. *Journal of Industrial and Production Engineering, 37*(2–3), 148–157. https://doi.org/10.1080/21681015.2020.1741038

Li, Y., Zhang, S., Zhang, X., & Wang, A. (2012). Agent-based partner selection and simulation in collaborative product development process. *Advances in Information Sciences and Service Sciences, 4*(20).

Lind, M., Visentini, M., Mäntylä, T., & Del Missier, F. (2017). Choice-supportive misremembering: A new taxonomy and review. *Frontiers in Psychology, 8*, 2062. https://doi.org/10.3389/fpsyg.2017.02062

Linton, J. D., Walsh, S. T., & Morabito, J. (2002). Analysis, ranking, and selection of R&D projects in a portfolio. *R&D Management, 32*(2), 139–148. https://doi.org/10.1111/1467-9310.00246

Liu, S., & Louis, S. J. (2016). Comparing two representations for evolving micro in 3D RTS games. In *2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 722–729). IEEE. https://doi.org/10.1109/ICTAI.2016.0114

Loch, C. (2000). Tailoring product development to strategy: Case of a European technology manufacturer. *European Management Journal, 18*(3), 246–258. https://doi.org/10.1016/s0263-2373(00)00007-4

Loch, C. H., & Kavadias, S. (2002). Dynamic portfolio selection of NPD programs using marginal returns. *Management Science, 48*(10), 1227–1241.

Loch, C. H., De Meyer, A., & Pich, M. T. (2006). *Managing the unknown: A new approach to managing high uncertainty and risk in projects*. John Wiley & Sons.

Lyons, D., Finocchiaro, J., Novitzky, M., & Korpela, C. (2020). A Monte Carlo approach to closing the reality gap. *arXiv preprint arXiv:2005.03809.*

MacMillan, I. C., & McGrath, R. G. (2002). Crafting R&D project portfolios. *Research-Technology Management, 45*(5), 48–59. https://doi.org/10.1080/08956308.2002.11671522

Malcolm, D., Roseboom, J., Clark, C., & Fazar, W. (1959). Application of a technique for research and development program evaluation. *Operations Research, 7*(5), 646–669. https://doi.org/10.1287/opre.7.5.646

Marcondes, G. (2021). Multicriteria decision method for project ranking considering uncertainty. In *Proceedings of the 10th International Conference on Operations Research and Enterprise Systems* (pp. 123–128). SCITEPRESS. https://doi.org/10.5220/0010183601230128

Marcondes, G. A. B., & Vilela, M. da S. (2022). Project selection with uncertainty using Monte Carlo simulation and multi-criteria decision methods. In *Operations Research and Enterprise Systems* (pp. 152–170). Springer. https://doi.org/10.1007/978-3-031-10725-2_8

Marcondes, G., & Marcondes, C. (2024). Innovation project selection considering stochastic weighted product model. In *Proceedings of the 13th International Conference on Operations Research and Enterprise Systems* (pp. 207–212). SCITEPRESS. https://doi.org/10.5220/0012270000003639

Mardani, A., Jusoh, A., & Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications – Two decades review from 1994 to 2014. *Expert Systems with Applications, 42*(8), 4126–4148.

Margetts, T. J., Karnik, S. J., Wang, H. S., Plotkin, L. I., Oblak, A. L., Fehrenbacher, J. C., ... & Movila, A. (2024). Use of AI language engine ChatGPT 4.0 to write a scientific review article examining the intersection of Alzheimer's disease and bone. *Current Osteoporosis Reports, 22*(1), 177–181.

Martinsuo, M. (2013). Project portfolio management in practice and in context. *International Journal of Project Management, 31*(6), 794–803. https://doi.org/10.1016/j.ijproman.2012.10.013

Mavrotas, G., & Makryvelios, E. (2021). Combining multiple criteria analysis, mathematical programming and Monte Carlo simulation to tackle uncertainty in research and development project portfolio selection: A case study from Greece. *European Journal of Operational Research, 291*(2), 794–806. https://doi.org/10.1016/j.ejor.2020.09.051

Mavrotas, G., & Pechak, O. (2013). The trichotomic approach for dealing with uncertainty in project portfolio selection: Combining MCDA, mathematical programming and Monte Carlo simulation. *International Journal of Multicriteria Decision Making, 3*(1), 79–95. https://doi.org/10.1504/IJMCDM.2013.052474

McConnell, S. (2006). *Software estimation: Demystifying the black art*. Microsoft Press.

Medaglia, A. L., Graves, S. B., & Ringuest, J. L. (2007). A multiobjective evolutionary approach for linearly constrained project selection under uncertainty. *European Journal of Operational Research, 179*(3), 869–894. https://doi.org/10.1016/j.ejor.2005.03.068

Menke, M. M. (1997). Essentials of R&D strategic excellence. *Research-Technology Management, 40*(5), 42–47. https://doi.org/10.1080/08956308.1997.11671156

Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F. Å., & Lanamäki, A. (2015). "The sum of all human knowledge": A systematic review of scholarly research on the content of Wikipedia. *Journal of the Association for Information Science and Technology, 66*(2), 219–245. https://doi.org/10.1002/asi.23172

Mikkola, J. H. (2001). Portfolio management of R&D projects: Implications for innovation management. *Technovation, 21*(7), 423–435. https://doi.org/10.1016/s0166-4972(00)00062-6

Mir, R., & Greenwood, M. (2021). *Philosophy and management studies: A research overview*. Routledge.

Mourão, E., Pimentel, J. F., Murta, L., Kalinowski, M., Mendes, E., & Wohlin, C. (2020). On the performance of hybrid search strategies for systematic literature reviews in software engineering. *Information and Software Technology, 123*, 106294.

Mozelius, P., & Humble, N. (2024). On the use of generative AI for literature reviews: An exploration of tools and techniques. In *Proceedings of the 23rd European Conference on Research Methodology for Business and Management Studies*. Academic Conferences and Publishing Limited.

Myers, S. C. (1977). Finance theory and financial strategy. In E. S. Schwartz & L. Trigeorgis (Eds.), *Real options and investment under uncertainty: Classical readings and recent contributions* (2001). https://doi.org/10.1287/inte.14.1.126

Naldi, M., Nicosia, G., Pacifici, A., & Pferschy, U. (2019). Profit-fairness trade-off in project selection. *Socio-Economic Planning Sciences, 67*, 133–146. https://doi.org/10.1016/j.seps.2018.10.007

Narayanan, M., Abdulazeez, M., Bukhamsin, K., & Alshehri, N. (2023). Stochastic economic ranking – A prudent way to address risk and uncertainty for decision-makers. *Middle East Oil, Gas and Geosciences Show*. https://doi.org/10.2118/213385-MS

Nazzal, M. K., Morris, A. J., Parker, R. S., White, F. A., Natoli, R. M., Fehrenbacher, J. C., & Kacena, M. A. (2024). Using AI to write a review article examining the role of the nervous system on skeletal homeostasis and fracture healing. *Current Osteoporosis Reports, 22*(1), 217–221.

Neely III, J. E., & De Neufville, R. (2001). Hybrid real options valuation of risky product development projects. *International Journal of Technology, Policy and Management, 1*(1), 29. https://doi.org/10.1504/ijtpm.2001.001743

Nelson, R. R. (1961). Uncertainty, learning, and the economics of parallel research and development efforts. *The Review of Economics and Statistics, 43*(4), 351–364. https://doi.org/10.2307/1927475

Nielsen, M. K., Jacobsen, A. M., Carstensen, J. L., Nielsen, M. T., & Tambo, T. (2024). Industrial R&D project portfolio selection method using a multi-objective optimisation program: A conceptual quantitative framework. *Journal of Industrial Engineering and Management, 17*(1), 217. https://doi.org/10.3926/jiem.6552

OpenAI. (2024). *Introducing ChatGPT*. OpenAI. https://openai.com/index/chatgpt

Pace, D. K. (2013). Comprehensive consideration of uncertainty in simulation use. *The Journal of Defense Modeling and Simulation, 10*(4), 367–380.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Panadero, J., Doering, J., Kizys, R., Juan, A. A., & Fito, A. (2020). A variable neighborhood search simheuristic for project portfolio selection under uncertainty. *Journal of Heuristics, 26*(3), 353–375. https://doi.org/10.1007/s10732-018-9367-z

Pendharkar, P. C. (2014). A decision-making framework for justifying a portfolio of IT projects. *International Journal of Project Management, 32*(4), 625–639. https://doi.org/10.1016/j.ijproman.2013.09.006

Perlitz, M., Peske, T., & Schrank, R. (1999). Real options valuation: The new frontier in R&D project evaluation? *R&D Management, 29*(3), 255–270. https://doi.org/10.1111/1467-9310.00135

Pillai, A. S., Joshi, A., & Rao, K. S. (2002). Performance measurement of R&D projects in a multi-project, concurrent engineering environment. *International Journal of Project Management, 20*(2), 165–177. https://doi.org/10.1016/s0263-7863(00)00056-9

Posen, H. E., Leiblein, M. J., & Chen, J. S. (2018). Toward a behavioral theory of real options: Noisy signals, bias, and learning. *Strategic Management Journal, 39*(4), 1112–1138.

Python Software Foundation. (2024). *Python*. Retrieved from https://www.python.org/

Qureshi, R., Shaughnessy, D., Gill, K. A., Robinson, K. A., Li, T., & Agai, E. (2023). Are ChatGPT and large language models “the answer” to bringing us closer to systematic review automation? *Systematic Reviews, 12*(1), 72.

Rianmora, S., & Werawatganon, S. (2021). Applying quality function deployment in open innovation engineering. *Journal of Open Innovation: Technology, Market, and Complexity, 7*(1), 26.

Roeth, T., Spieth, P., & Joachim, V. (2019). The interaction of intuition and rationality during escalated NPD decisions: An investigation of decision-makers' affective states. *International Journal of Innovation Management, 24*(4), 2050033. https://doi.org/10.1142/s1363919620500334

Ross, J., Fisch, J. H., & Varga, E. (2017). Unlocking the value of real options: How firm-specific learning conditions affect R&D investments under uncertainty. *Strategic Entrepreneurship Journal, 12*(3), 335–353. https://doi.org/10.1002/sej.1275

Roy, R., Ashmika, R., Chakraborty, A., & Sharafat, I. (2024). Future trends in AI and academic research writing. In *Utilizing AI tools in academic research writing* (pp. 232–254). IGI Global.

Sahlman, W. A. (1990). The structure and governance of venture-capital organisations. *Journal of Financial Economics, 27*(2), 473–521. https://doi.org/10.1016/0304-405X(90)90065-8

Saiz, M., Calvet, L., Juan, A. A., & Lopez-Lopez, D. (2024). A simheuristic for project portfolio optimization combining individual project risk, scheduling effects, interruptions, and project risk correlations. *Computers & Industrial Engineering, 198*, 110694. https://doi.org/10.1016/j.cie.2024.110694

Sarangee, K. R., Woolley, J. L., Schmidt, J. B., & Long, E. (2014). De-escalation mechanisms in high-technology product innovation. *Journal of Product Innovation Management, 31*(5), 1023–1038. https://doi.org/10.1111/jpim.12142

Saunders, M., Lewis, P., & Thornhill, A. (2022). *Research methods for business students* (9th ed.). Pearson.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Schmidt, J. B., Montoya-Weiss, M. M., & Massey, A. P. (2001). New product development decision-making effectiveness: Comparing individuals, face-to-face teams, and virtual teams. *Decision Sciences, 32*(4), 575–600. https://doi.org/10.1111/j.1540-5915.2001.tb00973.x

Schmidt, W. J., Rincon-Yanez, D., Kharlamov, E., & Paschke, A. (2022). Scaling scientific knowledge discovery with neuro-symbolic AI and large language models.

Schrader, S., Riggs, W. M., & Smith, R. P. (1993). Choice over uncertainty and ambiguity in technical problem solving. *Journal of Engineering and Technology Management, 10*(1–2), 73–99.

Schumpeter, J. A. (1934). *The theory of economic development* (R. Opie, Trans.). Harvard University Press.

Schwalbe, K. (2009). *Introduction to project management*. Course Technology Cengage Learning.

Schwartz, E., & Moon, M. (2000). Rational pricing of internet companies. *Financial Analysts Journal, 56*(3), 62–75. https://doi.org/10.2469/faj.v56.n3.2361

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews Neuroscience, 23*(7), 439–452. https://doi.org/10.1038/s41583-022-00587-4

Shakhsi-Niaei, M., Torabi, S. A., & Iranmanesh, S. H. (2011). A comprehensive framework for project selection problem under uncertainty and real-world constraints. *Computers & Industrial Engineering, 61*(1), 226–237. https://doi.org/10.1016/j.cie.2011.03.015

Shani, C., Vreeken, J., & Shahaf, D. (2023). Towards concept-aware large language models. *arXiv preprint arXiv:2311.01866.* https://arxiv.org/abs/2311.01866

Sheasley, W. D. (2000). Taking an options approach to new technology development. *Research-Technology Management, 43*(6), 37–43. https://doi.org/10.1080/08956308.2000.11671394

Sherman, S. J., & Corty, E. (1984). Cognitive heuristics. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 189–286). Lawrence Erlbaum Associates Publishers.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*(2), 129–138.

Simplício, J. G., Lemme, C. F., & Leal, R. P. C. (2012). Portfolio theory in the selection of oil investment projects. *Gestão & Produção, 19*(2), 265–272. https://doi.org/10.1590/s0104-530x2012000200003

SimPy. (2024). *SimPy – Discrete event simulation for Python*. Retrieved 2024, from https://simpy.readthedocs.io/en/latest/

Sina, L. B., Secco, C. A., Blazevic, M., & Nazemi, K. (2024). Guided visual analytics – A visual analytics guidance approach for systematic reviews in research. In *Artificial intelligence and visualization: Advancing visual knowledge discovery* (pp. 319–343). Springer Nature Switzerland.

Singh, H., Cascini, G., & McComb, C. (2022). Idea selection in design teams: A computational framework and insights in the presence of influencers. *Design Science, 8*, e17. https://doi.org/10.1017/dsj.2022.17

Smit, H. T. J., & Trigeorgis, L. (2007). Strategic options and games in analysing dynamic technology investments. *Long Range Planning, 40*(1), 84–114. https://doi.org/10.1016/j.lrp.2007.02.005

Snowden, D. (2005). Strategy in the context of uncertainty. *Handbook of Business Strategy, 6*(1), 47–54. https://doi.org/10.1108/08944310510556955

Sommer, S. C., Loch, C. H., & Dong, J. (2009). Managing complexity and unforeseeable uncertainty in startup companies: An empirical study. *Organization Science, 20*(1), 118–133. https://doi.org/10.1287/orsc.1080.0369

Souifi, L., Khabou, N., Rodriguez, I. B., & Kacem, A. H. (2024). Towards the use of AI-based tools for systematic literature review. *ICAART (2)*, 595–603.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organisational Behavior and Human Performance, 16*(1), 27–44. https://doi.org/10.1016/0030-5073(76)90005-2

Steyvers, M., & Kumar, A. (2023). Three challenges for AI-assisted decision-making. *Perspectives on Psychological Science, 19*(1), 17456916231181102. https://doi.org/10.1177/17456916231181102

Sulis, E., & Taveter, K. (2022). *Agent-based business process simulation*. Springer, Cham.

Sun, P., & Zuo, X. (2024). Philosophical foundations of management research: A comprehensive review. *Journal of Scientific Reports, 6*(1), 1–22.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509–533.

Teglio, A., Raberto, M., & Cincotti, S. (2009). Explaining equity excess return by means of an agent-based financial market. In *Artificial Economics* (pp. 145–156). https://doi.org/10.1007/978-3-642-02956-1_12

Teirlinck, P., & Spithoven, A. (2012). Formal R&D management and strategic decision-making in small firms in knowledge-intensive business services. *R&D Management, 43*(1), 37–51. https://doi.org/10.1111/j.1467-9310.2012.00701.x

Tinoco, M. A. C., Dutra, C. C., Ribeiro, J. L. D., Miorando, R. F., & Caten, C. S. T. (2018). An integrated model for evaluation and optimisation of business project portfolios. *European Journal of Industrial Engineering, 12*(3), 442–465. https://doi.org/10.1504/EJIE.2018.092010

Tomczyk, P., Brüggemann, P., Mergner, N., & Petrescu, M. (2024). Exploring AI's role in literature searching: Traditional methods versus AI-based tools in analyzing topical e-commerce themes. In *Digital Marketing & eCommerce Conference* (pp. 141–148). Springer, Cham.

Tóth, B., & Oldal, L. G. (2022). Automation of systematic literature reviews: Development of a gold standard of articles in PubMed 2020–2021, a research protocol. In *2022 IEEE 20th Jubilee International Symposium on Intelligent Systems and Informatics (SISY)* (pp. 59–64). IEEE.

Trigeorgis, L. (1996). *Real options: Managerial flexibility and strategy in resource allocation*. MIT Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

U.S. Bureau of Labor Statistics. (2024, January 11). *1-year survival rates for new business establishments, by year and location*. TED: The Economics Daily. https://www.bls.gov/opub/ted/2024/1-year-survival-rates-for-new-business-establishments-by-year-and-location.htm

University of Waikato. (n.d.). *Guidelines for student use of generative AI tools*. University of Waikato. Retrieved February 23, 2025, from https://www.waikato.ac.nz/assets/Uploads/Student-life/Student-assessment-Handbook/Guidelines-for-student-use-of-generative-AI-tools-FINAL.pdf

Vaculik, M., Lorenz, A., Roijakkers, N., & Vanhaverbeke, W. (2019). Pulling the plug? Investigating firm-level drivers of innovation project termination. *IEEE Transactions on Engineering Management, 66*(2), 180–192. https://doi.org/10.1109/tem.2018.2798922

Vihari, N. S., & Kaur, A. (2024). The role of generative AI-assisted literature reviews in transforming academic research. In *Utilizing AI tools in academic research writing* (pp. 77–87). IGI Global.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology, 37*(2), 209–226.

Wang, J., Lin, W., & Huang, Y. H. (2010). A performance-oriented risk management framework for innovative R&D projects. *Technovation, 30*(11–12), 601–611. https://doi.org/10.1016/j.technovation.2010.07.003

Whelan, J., Ghoniem, M., Médoc, N., Apicella, M., & Beck, E. (2021). Applying a novel approach to scoping review incorporating artificial intelligence: Mapping the natural history of gonorrhoea. *BMC Medical Research Methodology, 21*(1). https://doi.org/10.1186/s12874-021-01367-x

Wieber, F., Thürmer, J. L., & Gollwitzer, P. M. (2015). Attenuating the escalation of commitment to a faltering project in decision-making groups: An implementation intention approach. *Social Psychological and Personality Science, 6*(5), 587–595.

Wikipedia. (2024). *List of academic databases and search engines*. Retrieved February 1, 2024, from https://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines

Woolley, A. W., & Gupta, P. (2024). Understanding collective intelligence: Investigating the role of collective memory, attention, and reasoning processes. *Perspectives on Psychological Science, 19*(2), 344–354.

Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective intelligence in teams and organisations. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence* (pp. 143–168). MIT Press.

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research, 39*(1), 93–112.

Xuan, H., Yao, C., Li, H., & Chang, X. (2021). Portfolio model based on scenario tree. *E3S Web of Conferences, 251*, 01114. https://doi.org/10.1051/e3sconf/202125101114

Yang, M. M., Chen, W., & Wang, Y. (2020). What explains managers' escalating behaviors in a failing NPD project? The impact of managerial perceptions of opportunities and threats in a stage-gate process. *Journal of Small Business Management, 60*(3), 541–579. https://doi.org/10.1080/00472778.2020.1719297

Yao, X., Kumar, M. V., Su, E., Flores Miranda, A., Saha, A., & Sussman, J. (2024). Evaluating the efficacy of artificial intelligence tools for the automation of systematic reviews in cancer research: A systematic review. *Cancer Epidemiology, 88*, 102511. https://doi.org/10.1016/j.canep.2023.102511

Zala, K., Acharya, B., Mashru, M., Palaniappan, D., Gerogiannis, V. C., Kanavos, A., & Karamitsos, I. (2024). Transformative automation: AI in scientific literature reviews. *International Journal of Advanced Computer Science and Applications, 15*(1). https://doi.org/10.14569/ijacsa.2024.01501122

Zaman, M., & Tanewski, G. (2024). R&D investment, innovation, and export performance: An analysis of SME and large firms. *Journal of Small Business Management*, 1–34.

Zammar, A., Luiz Kovaleski, J., & Negri Pagani, R. (2023). Innovation management tools: A comprehensive literature approach of the last three decades. *Management Review Quarterly.* https://doi.org/10.1007/s11301-023-00329-5

Zeigler, B. P., Muzy, A., & Kofman, E. (2018). *Theory of modeling and simulation: Discrete event & iterative system computational foundations* (3rd ed.). Academic Press.

Appendix A - Source Selection and Journal List

Figure A1

Scopus Source Selection Tool

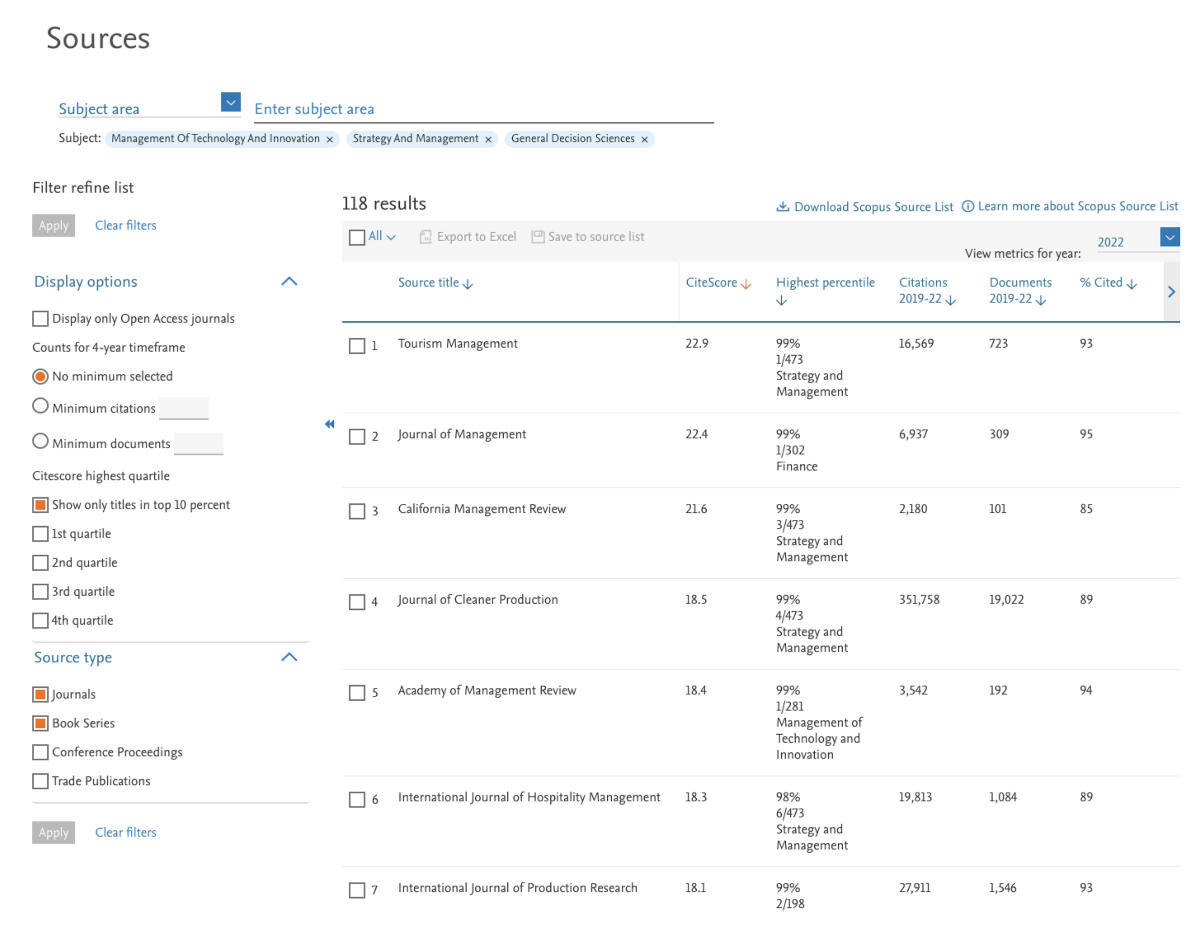


Table A1

Journal List

Relevance	Citescore	ABDC[5]	Title
5	99%	A*	Journal of Business Venturing
5	96%	A*	Journal of Product Innovation Management
5	93%	A*	Strategic Management Journal
5	91%	A*	Decision Sciences
5	95%	A	International Journal of Project Management
5	94%	A	Academy of Management Perspectives
5	94%	A	Global Strategy Journal
5	91%	A	Strategic Entrepreneurship Journal
5	98%	-	Journal of Innovation and Knowledge
5	96%	-	Decision Making: Applications in Management and Engineering
5	94%	-	R and D Management
5	92%	-	Journal of Innovation and Entrepreneurship
4	99%	A*	Journal of Management
4	99%	A*	Academy of Management Review
4	98%	A*	Academy of Management Journal
4	94%	A*	Journal of Management Studies
4	99%	A	California Management Review
4	98%	A	Technological Forecasting and Social Change
4	98%	A	Long Range Planning
4	97%	A	Asia Pacific Journal of Management
4	96%	A	Technovation
4	96%	A	Journal of Small Business Management
4	95%	A	International Journal of Management Reviews
4	95%	A	Strategic Organization
4	94%	A	British Journal of Management
4	92%	A	Journal of International Management
4	95%	B	Journal of Leadership and Organizational Studies
4	94%	B	Economics of Innovation and New Technology
4	93%	B	Management Review Quarterly
4	92%	B	European Management Journal
4	96%	-	Cross Cultural and Strategic Management
4	95%	-	Entrepreneurial Business and Economics Review
4	94%	-	European Research on Management and Business Economics
4	92%	-	Journal of Engineering and Technology Management - JET-M
4	92%	-	World Journal of Entrepreneurship, Management and Sustainable Development
4	90%	-	Business Strategy and Development

Where *<JOURNAL LIST>* is referred to in Scopus searches, it should be substituted with:

> SRCTITLE(Journal of Business Venturing) OR SRCTITLE(Journal of Product Innovation Management) OR SRCTITLE(Strategic Management Journal) OR SRCTITLE(Decision Sciences) OR SRCTITLE(International Journal of Project Management) OR SRCTITLE(Global Strategy Journal) OR SRCTITLE(Strategic

[5] Australian Business Deans Council, 2023

Entrepreneurship Journal) OR SRCTITLE(Journal of Innovation and Knowledge) OR SRCTITLE(Decision Making: Applications in Management and Engineering) OR SRCTITLE(R and D Management) OR SRCTITLE(Journal of Innovation and Entrepreneurship) OR SRCTITLE(Journal of Management) OR SRCTITLE(Academy of Management Review) OR SRCTITLE(Journal of International Business Studies) OR SRCTITLE(Academy of Management Journal) OR SRCTITLE(Journal of Management Studies) OR SRCTITLE(California Management Review) OR SRCTITLE(Technological Forecasting and Social Change) OR SRCTITLE(Long Range Planning) OR SRCTITLE(Asia Pacific Journal of Management) OR SRCTITLE(Technovation) OR SRCTITLE(Journal of Small Business Management) OR SRCTITLE(International Journal of Management Reviews) OR SRCTITLE(Strategic Organization) OR SRCTITLE(Academy of Management Perspectives) OR SRCTITLE(British Journal of Management) OR SRCTITLE(Journal of International Management) OR SRCTITLE(Journal of Leadership and Organizational Studies) OR SRCTITLE(Economics of Innovation and New Technology) OR SRCTITLE(Management Review Quarterly) OR SRCTITLE(European Management Journal) OR SRCTITLE(Cross Cultural and Strategic Management) OR SRCTITLE(Entrepreneurial Business and Economics Review) OR SRCTITLE(European Research on Management and Business Economics) OR SRCTITLE(Journal of Engineering and Technology Management - JET-M) OR SRCTITLE(World Journal of Entrepreneurship, Management and Sustainable Development) OR SRCTITLE(Business Strategy and Development)

Appendix B – Exploratory Searches

Table B1 shows the initial sequence of keyword searches tested against Scopus (using the online user interface, not developed software). The objective was to identify keywords that limited the result pool so that it could be processed programmatically.

Table B1

Preliminary Scopus Searches

TITLE-ABS-KEY	PUBDATE	Results
(all selected journal articles)	≥ 2010	242631
research and development	≥ 2010	23353
decision making	≥ 2010	17130
innovation management	≥ 2010	8607
product management	≥ 2010	8175
new product development	≥ 2010	3072
portfolio management	≥ 2010	1521
research and development, governance	≥ 2010	1054
decision making, product management	≥ 2010	795
decision making, innovation management	≥ 2010	591
real options	≥ 2010	578
research and development, portfolio	≥ 2010	327
research and development, decision making, product management	≥ 2010	173
venture governance	≥ 2010	166
research and development, decision making, options	≥ 2010	87
research and development, product, portfolio management	≥ 2010	73
decision making, innovation management, uncertainty	≥ 2010	63
research and development, decision making, entrepreneur	≥ 2010	62
research and development, governance, new product development	≥ 2010	41
research and development, venture governance	≥ 2010	27
research and development, product, portfolio management, decision making	≥ 2010	12

Notes:

1. Scopus does not differentiate between “decision making” and “decision-making”.
2. Keywords (TITLE-ABS-KEY) are combined using boolean AND.
3. The journal list in Appendix A is used in all Scopus searches as SRCTITLE parameters.
4. ***Highlighted italic*** items were carried forward into the final search design.

During the development of the software, it was necessary to test various searches to explore its capabilities and limitations and to refine the search keywords themselves. Table B2 captures the searches that were used for this purpose. Note that these did not contribute to the final result set.

Table B2

Exploratory Searches

#	Time of Search	Earliest Year	Search keywords	Results
1	06-Apr-24 18:16:25	2010	"research and development" AND "decision making" AND "options"	87
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(decision making) AND TITLE-ABS-KEY(options)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
2	06-Apr-24 18:27:43	2010	"research and development" AND "decision making" AND "product management"	173
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(decision making) AND TITLE-ABS-KEY(product management)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
3	07-Apr-24 10:09:25	2010	"decision making" AND "innovation management" AND "uncertainty"	63
	(TITLE-ABS-KEY(decision making) AND TITLE-ABS-KEY(innovation management) AND TITLE-ABS-KEY(uncertainty)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
4	07-Apr-24 10:14:06	2010	"research and development" AND "venture governance"	27
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(venture governance)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
5	07-Apr-24 10:16:41	2010	"research and development" AND "governance" AND "new product development"	41
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(governance) AND TITLE-ABS-KEY(new product development)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
6	07-Apr-24 10:20:49	2010	"research and development" AND "decision making" AND "entrepreneur"	62
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(decision making) AND TITLE-ABS-KEY(entrepreneur)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
7	07-Apr-24 10:29:18	2010	"research and development" AND "product" AND "portfolio management"	73

	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(product) AND TITLE-ABS-KEY(portfolio management)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
8	07-Apr-24 15:37:40	2010	research and development AND "product" AND "portfolio management" AND "decision making"	12
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(product) AND TITLE-ABS-KEY(portfolio management) AND TITLE-ABS-KEY(decision making)) AND(<JOURNAL LIST>) AND (PUBYEAR > 2009)			
9	07-Apr-24 15:45:30	1900	research and development AND "product" AND "portfolio management" AND "decision making"	22
	(TITLE-ABS-KEY(research and development) AND TITLE-ABS-KEY(product) AND TITLE-ABS-KEY(portfolio management) AND TITLE-ABS-KEY(decision making)) AND(<JOURNAL LIST>) AND (PUBYEAR > 1899)			
10	07-Apr-24 15:46:39	1900	product AND "portfolio management" AND "decision making"	56
	(TITLE-ABS-KEY(product) AND TITLE-ABS-KEY(portfolio management) AND TITLE-ABS-KEY(decision making)) AND(<JOURNAL LIST>) AND (PUBYEAR > 1899)			

Appendix C – Selected Articles and Extracted Data[6]

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Behrens & Ernst (2013)	10.1111/jpim.12100	Quantitative	Hybrid/mixed approach	Unspecified	Unspecified	Visual indicators, Highlighting sunk costs, External consultant	Sunk cost fallacy, Management bias	2014
Boulding et al. (1997)	10.1177/002224379703400114	Quantitative	Stage-gate, New decision maker	Gatekeeper independence	NPV, Uncertainty, Opportunity cost	Unspecified	Commitment bias, Impression management	1997
Chan et al. (2007)	10.1287/mnsc.1060.0676	Quantitative	Unspecified	Organisational agility, Dynamic capabilities	Transaction cost, Project interdependency, Market value, Costs	Simulation, Dynamic programming	Unspecified	2007
Cheung et al. (2009)	10.1177/030630700903500205	Qualitative	Satisficing, Formalised process	Unspecified	Unspecified	Process evaluation tool	Unspecified	2009
Cooper (2006)	10.1080/08956308.2006.11657405	Qualitative	Stage-gate, Portfolio management	Defined strategy, Budget allocation	Strategic alignment, Probability of technical success, Probability of commercial success, Benefits	Scoring model, Force-ranking	Risk aversion, Misuse of NPV	2006

[6] Available at https://docs.google.com/spreadsheets/d/1pYN9Ml-KUgiBBqViq9Q1Q4Yy74dm4wSAvlMgpo-uY8M

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Cooper et al. (2000)	10.1080/08956308.2000.11671338	Qualitative	Portfolio management, Formalised process	Precommitted rule, Defined strategy	Capacity, Capability, Strategic alignment, Balanced portfolio, Project ranking, Economic capital value	Force-ranking, Resource-capacity analysis, Check-lists, Scoring model	Inadequate data, Too many projects, Misuse of NPV	2000
Eliens et al. (2018)	10.1111/jpim.12452	Quantitative	Hybrid/mixed approach, Formalised process	Gatekeeper independence	Unspecified	Unspecified	Self-justification, Belief inertia, Loss aversion, Endowment effect, Status quo bias, Confirmation bias	2018
Hauser & Zettelmeyer (1997)	10.1080/08956308.1997.11671140	Mixed	Hybrid/mixed approach, Real options, Portfolio management	Unspecified	Benefits, Costs, Strategic alignment, Balanced portfolio	Unspecified	Not invented here, Risk aversion	1997
Jagle (1999)	10.1111/1467-9310.00136	Mixed	Real options	Unspecified	NPV, Risks, Costs, Benefits, Market value, Costs	Decision tree, Sensitivity analysis	Misuse of NPV	1999
Linton et al. (2002)	10.1111/1467-9310.00246	Mixed	MCDM, Data envelopment analysis	Unspecified	Unspecified	Visual indicators	Unspecified	2002

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Loch (2000)	10.1016/S0263-2373(00)00007-4	Mixed	Hybrid/mixed approach, Formalised process	Customer orientation, Project champion, Project manager, Market awareness, Defined strategy	Management support	Unspecified	Unspecified	2000
MacMillan & McGrath (2002)	10.1080/08956308.2002.11671522	Quantitative	Real options, Portfolio management	Unspecified, Defined strategy	Market value, NPV, Supply chain, Technical feasibility, Competition, Pace of technological change, Capacity, Capability, Risks, Budget, Management support, Benefits, Costs, Balanced portfolio	Resource-capacity analysis	Underinvesting in future, Too many projects	2002

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Menke (1997)	10.1080/08956308.1997.11671156	Quantitative	Hybrid/mixed approach	Market awareness, Customer orientation, Defined strategy, Project manager, Project tracking, Cross-functional teams, Coordinated GTM, Clear requirements	Management support, Sufficient resources	Unspecified	Unspecified	1997
Mikkola (2001)	10.1016/S0166-4972(00)00062-6	Qualitative	Formalised process	Unspecified	Competitive Advantage, Customer benefits	Selection matrix	Unspecified	2001
Neely III & De Neufville (2001)	10.1504/ijtpm.2001.001743	Quantitative	Real options, Decision analysis	Unspecified	Uncertainty, Costs, Benefits, Comparable asset valuation	Sensitivity analysis, Decision tree	Estimation error, Market volatility	2001
Nielsen et al. (2024)	10.3926/jiem.6552	Quantitative	Portfolio management	Innovation strategy, Market awareness, Defined risk orientation, Budget allocation	Constraints, MOOP, NPV, Strategic alignment, Disruptive rank, Capacity, Capability, Benefits	NPV	Uncertainty, Management bias	2024

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Perlitz et al. (1999)	10.1111/1467-9310.00135	Qualitative	Real options	Unspecified	NPV, Uncertainty, Budget, Risks, Benefits, Market value	Unspecified	Unspecified	1999
Pillai et al. (2002)	10.1016/S0263-7863(00)00056-9	Qualitative	Portfolio management, Decision effectiveness	Continuous forecasting, Stakeholder engagement	NPV, Uncertainty, Assumption validity	Scoring model, Graphical cost-scope-time, Selection Matrix	Market volatility, No longer a strategic fit	2002
Roeth et al. (2019)	10.1142/S1363919620500334	Quantitative	Hybrid/mixed approach	Positive affect, Intuitive DMS, Gatekeeper independence	Unspecified	Unspecified	Negative affect, Sunk cost fallacy	2020
Ross et al. (2017)	10.1002/sej.1275	Quantitative	Learning first, Real options	Unspecified	Unspecified	Unspecified	Uncertainty	2018
Sarangee et al. (2014)	10.1111/jpim.12142	Qualitative	Hybrid/mixed approach, Formalised process	Gatekeeper independence, Strategic options, Failure tolerance, Project tracking, Precommitted rule, Clear requirements, Market awareness	Unspecified, Assumption validity	Benchmarking, Roadmaps, Comprehensive testing	Reporting bias, Management bias	2014
Schmidt et al. (2001)	10.1111/j.1540-5915.2001.tb00973.x	Quantitative	Group	Unspecified	Unspecified	Low-fidelity communication	Social pressure, Appeal to authority	2001

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Sheasley (2000)	10.1080/08956308.2000.11671394	Qualitative	Real options, Stage-gate, Portfolio management, Learning first	Precommitted rule	Market value, Uncertainty, Management support	Scoring model	Unspecified	2000
Smit & Trigeorgis (2007)	10.1016/j.lrp.2007.02.005	Mixed	Real options	Market awareness	Shareholder value	Simulation, Sensitivity analysis, Game theory	Estimation error, Market volatility	2007
Vaculik et al. (2019)	10.1109/TEM.2018.2798922	Quantitative	MCDM, Stage-gate, Formalised process	Innovation culture, Innovation strategy, Organisational agility, Market internationalisation, R&D department, Market awareness, Large firm size, Innovation experience	Capacity, Capability, Management support, Budget, Risks, NPV, Opportunity cost, Uncertainty, Ambiguity, Strategic alignment, Technical feasibility	Unspecified	Inadequate data, Cognitive biases, Over-reliance on past data	2019
Wang et al. (2010)	10.1016/j.technovation.2010.07.003	Qualitative	Quality function deployment	Unspecified	Unspecified	Scoring model, House of Quality, Balanced scorecard	Unspecified	2010

Author	DOI	Methodology	Approach	Antecedent	Criterion	Tool	Error Source	PubYear
Yang et al. (2020)	10.1080/00472778.2020.1719297	Quantitative	Stage-gate	Precommitted rule	Critical threat, Performance threat	Unspecified	Gates with no teeth, Sunk cost fallacy, Self-justification	2022
Zammar et al. (2023)	10.1007/s11301-023-00329-5	Qualitative	MCDM	Unspecified	Unspecified	Unspecified	Unspecified	2023

Appendix D – Synthesis Attempts

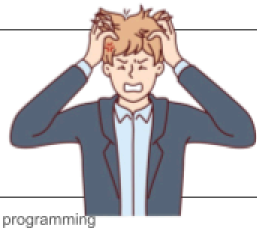
Figure D1

Synthesis – First Cut

	Start-up	SME	Large	Enterprise
Approach	Real options Decision analysis Strategic Bounded	Stage-Gate Learning-First Group Behavioural	Rational/Analytic Game theory Dynamic capabilities Decision effectiveness Data envelopment analysis	Selection Matrix Portfolio managemen Balanced scorecard MCDM Quality function deployment
Antecedents	Intuitive DMS Revalidate assumptions Customer orientation Defined strategy Market opportunity Innovation experience Positive affect Risk orientation Gatekeeper independence Having options Stakeholder engagement Financial capacity Precommitted rule Sufficient resources External consultant	Coordinate GTM Formalised process Project tracking Continuous forecasting Innovation strategy Organisational agility Rewarding good process R&D department Project manager Project champion Market awareness Budget allocation Tailored process	Failure tolerance Innovation culture Cross-functional teams Management support	Portfolio management Market internationalisation Large firm size Clear requirements Comparable asset valuation
Criteria	Budget Costs Market value Shareholder value Transaction cost Risks Benefits Customer benefits	NPV Project costs Capacity Ambiguity Uncertainty Competition Capability Constraints Competitive advantage Project interdependency	Management support Probability of commercial success Probability of technical success Technical feasibility	Balanced portfolio Supply chain Opportunity cost Pace of technological change Disruptive rank Project ranking Strategic alignment
Tools	MOOP New decision maker Assumption correction	NPV Decision tree ECV Feature-level de-escalation Resource-capacity analysis Check-lists Stage-Gate Comprehensive testing Roadmaps Process evaluation tool Project tracking Visual indicators Graphical cost-scope-time Highlighting sunk costs	Dynamic programming Internal competitions Low-fidelity communication	House of Quality Force-ranking Scoring model Selection matrix Benchmarking Step 1

Figure D2

Synthesis – Organisation by Discipline

	Start-up	Project Management	Innovation Culture	Portfolio Management
DM Approach – things we do	+Founder-led; Real options; Precommitted rule; Decision analysis; Assumption correction; +Lean startup; +Opportunism	Stage-Gate; Rational/Analytic; Decision effectiveness; +Board-led; Precommitted rule	Dynamic capabilities; +SLT-led; Game theory; Learning-First; Group; Behavioural	MCDM; Quality function deployment; Strategic alignment
Antecedents – things we have/are	Innovation experience; Having options; Customer orientation; Stakeholder engagement; Defined strategy; Bounded; Market opportunity; Sufficient resources	Formalised process; Project manager; +Budget tracking; Project champion; Project tracking; Continuous forecasting; Gatekeeper independence; Tailored process; Market awareness	Failure tolerance; Innovation strategy; Cross-functional teams; Management support; Risk orientation; Organisational agility; Positive affect; Rewarding good process; Intuitive DMS; R&D department	Portfolio management; Market internationalisation; Large firm size; Comprehensive testing
Criteria – things we confirm	Budget; Costs; Market value; Shareholder value; Risks; Customer benefits; Financial capacity; Technical feasibility	Project costs; Capacity; Ambiguity; Uncertainty; Competition; Capability; Constraints; Competitive advantage; Project interdependency; Clear requirements		Balanced portfolio; Supply chain; Opportunity cost; Pace of technological change; Disruptive rank
Tools – things we use	External consultant; ECV – expected commercial value; MOOP – maximum out of pocket	NPV; Decision tree; Feature-level de-escalation; Resource-capacity analysis; Project tracking; Stage-Gate; Process evaluation tool; Roadmaps; Check-lists; Visual indicators; Graphical cost-scope-time; Highlighting sunk costs	Dynamic programming; Internal competitions; Low-fidelity communication; +Design thinking	House of Quality; Force-ranking; Scoring model; Project ranking; Selection matrix; Benchmarking; Comparable asset valuation; Sensitivity analysis; Simulation

Step 2

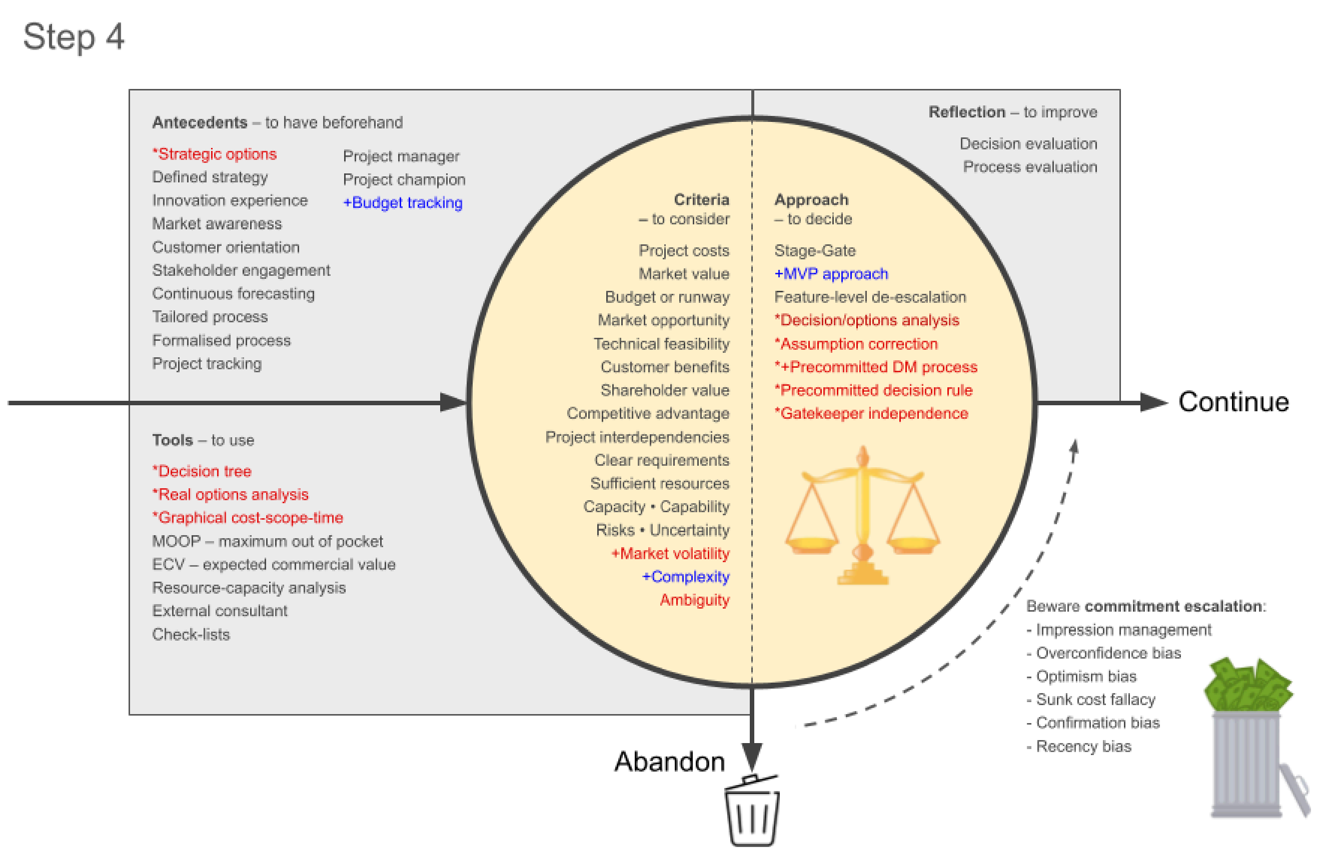
Figure D3

Synthesis – Refinement and Selection

	Start-up Founder-led, directive, intuitive *Entrepreneurial – opportunity as direction*		**Project (Innovation) Management** Governance-led, directive, rational *Planned – strategy as direction*		**Innovation Culture** SLT-led, facilitative, balanced *Exploratory – learning as direction*		**Portfolio (Product) Manageme** Discipline-led, facilitative, balanc *Stable - balance as direction*	
Approach how to decide	Decision/options analysis	*Precommitted rule *Assumption correction	+MVP approach Feature-level de-escalation Stage-Gate	*+Precommitted decision-making process *Precommitted go/kill criteria	Group DM	Learning-first Game theory	Portfolio Strategic alignment Quality function deployment	MCDM – multi-criteria (making
Criteria things to consider	Budget or runway Sufficient resources Market value Market opportunity +Biases	Risks Technical feasibility Shareholder value Customer benefits	Capacity Capability Uncertainty +Complexity *Ambiguity Clear requirements	Project costs Competitive advantage Project interdependencies *+Market volatility	+Climate +Culture +Resilience +Org structure +Engagement		Balanced portfolio Opportunity cost Disruptive rank	Pace of technological
Tools things to use	MOOP – maximum out of pocket NPV, ECV – expected commercial value	External consultant *Decision tree	Project tracking Graphical cost-scope-time Resource-capacity analysis	Check-lists Process evaluation Decision evaluation *Real options analysis	Dynamic programming Internal competitions +Design thinking +DELPHI +Agile	Visual indicators Highlighting sunk costs Low-fidelity communication	House of Quality Force-ranking Scoring model Project ranking Selection matrix Roadmaps	Benchmarkin Comparable a valuation Sensitivity an Simulation
Antecedents things to be or have	Innovation experience Market awareness Customer orientation	*Having strategic options Stakeholder engagement	Defined strategy Tailored process Project tracking Continuous forecasting Formalised process	*Gatekeeper independence Project manager +Budget tracking Project champion	Innovation strategy R&D department Cross-functional teams Management support Positive affect Failure tolerance	Organisational agility Process orientation Intuitive decision making system +Learning model Dynamic capabilities Risk orientation	Portfolio management Market internationalisation	Large firm siz Comprehensi testing Supply chain
Challenges things that limit decision making or cause errors	Bounded rationality +Optimism bias +Availability heuristic +Confirmation bias +Recency bias	+Negligible research budget + Overconfidence bias +Sunk cost fallacy +Impression management	+Limited research budget +Measurement noise +Anchoring bias	+Conflicting opinions Selected: Most relevant to SMEs	+Groupthink +Framing effect	+Status quo bias	+Loss aversion Step 3	

Figure D4

Synthesis – Initial Conceptual Model (Further Refinements Apply)



Appendix E – Simulation Platform

A lightweight simulation platform written in Google App Scripts (Google, 2024) was developed. This allowed close integration with Google Sheets and convenient access to graphing functions. The platform runs the following pseudocode and links to the spreadsheet as shown in Figure E1:

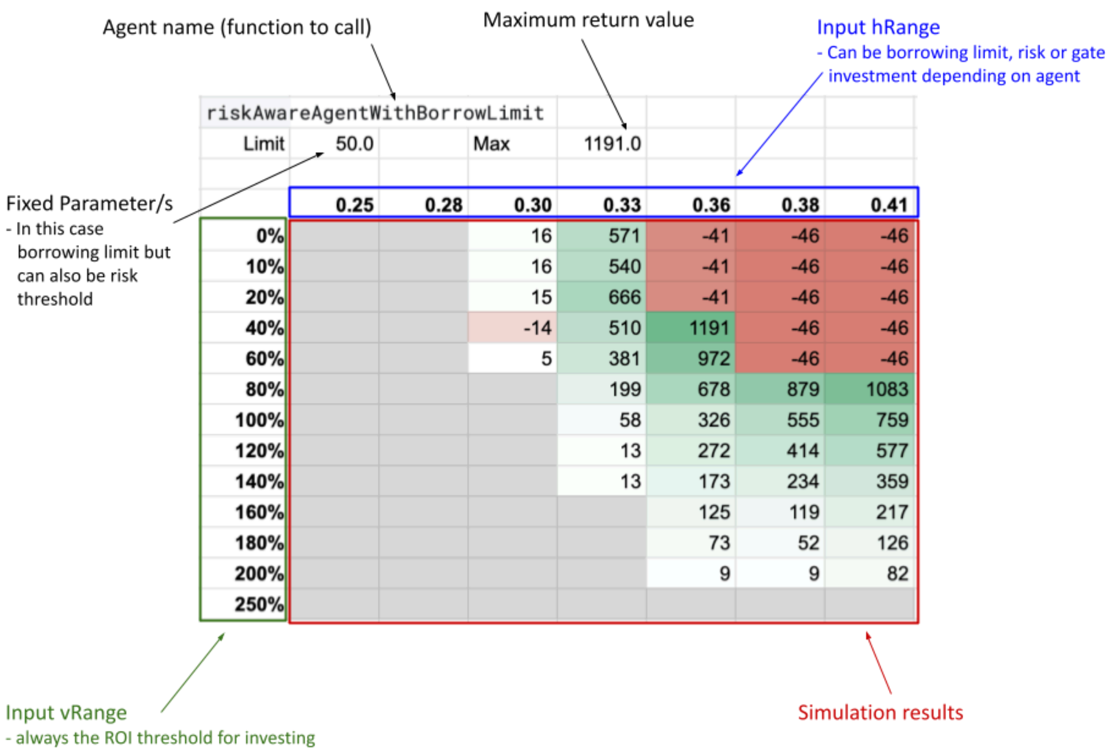
For each row in vRange (the ROI):

For each column in hRange:

Run the agent across the whole portfolio and capture the result in the respective cell.

Figure E1

Simulation Platform



	0.25	0.28	0.30	0.33	0.36	0.38	0.41
0%			16	571	-41	-46	-46
10%			16	540	-41	-46	-46
20%			15	666	-41	-46	-46
40%			-14	510	1191	-46	-46
60%			5	381	972	-46	-46
80%				199	678	879	1083
100%				58	326	555	759
120%				13	272	414	577
140%				13	173	234	359
160%					125	119	217
180%					73	52	126
200%					9	9	82
250%							

Licence

```
/**
* R&D Agent Simulation (this appendix) © 2024 by Owen Woollaston is
licensed under Creative Commons Attribution-NonCommercial 4.0
International. To view a copy of this license, visit
https://creativecommons.org/licenses/by-nc/4.0/
*/
```

Code – Simulation Platform

```
/**
* Simulates a process for an agent over a 2D grid of parameters.
*
* @param {string} agentName - The name of the agent being simulated.
* @param {...any} arguments - A flexible number of arguments, including grid
parameters.
* The function expects at least two arguments that are 2D arrays: one for the
vertical range and one for the horizontal range.
*
* 1. It extracts the provided arguments, and identifies two arrays:
*    - Vertical range (`vRange`) for rows.
*    - Horizontal range (`hRange`) for columns.
* 2. It checks if both ranges are provided, throwing an error if either is missing.
* 3. It loops through each combination of vertical and horizontal range values
(grid cells).
* 4. For each combination, it updates the argument array with the current row and
column values,
*    then calls the `runSim` function, passing the agent name and the modified
argument list.
* 5. The result of each simulation is stored in a 2D array (rows and columns),
which is then returned.
*
* Error Handling:
* - Throws an error if either the vertical or horizontal range is not properly
defined.
*
* @returns {Array} A 2D array of simulation results for each grid cell.
*
* Comment generated by ChatGPT. Code written by author.
*/
function gridSim(agentName) {
 var args = [...arguments].slice(1,arguments.length);

 // find the range arrays
 let hRange = 0;
 let vRange = 0;
 for(let i=2; i<args.length; i++) {
   if(Array.isArray(args[i])) {
     const [rows, cols] = [ args[i].length, args[i][0].length ];
     if(rows > 1 ) {
       vRange = i;
```

```
    } else if(cols && cols > 1) {
      hRange = i;
    }
  }
}

if( vRange==0 ) {
  throw "ERROR: You must define a vertical range";
}

if( hRange==0 ) {
  throw "ERROR: You must define a horizontal range";
}

var callArgs = args;
const hArgs = args[hRange][0];
const vArgs = args[vRange];
var rtn = [];
for(let r=0; r<vArgs.length; r++) {
  let row = [];
  for(let c=0; c<hArgs.length; c++) {
    callArgs[vRange] = vArgs[r][0];
    callArgs[hRange] = hArgs[c];
    let v = runSim(agentName, ...callArgs);
    row.push(v);
  }
  rtn.push(row);
}
return rtn;
}


/**
* Runs a simulation for an agent using effort and revenue arrays.
*
* @param {string} agentName - The name of the agent (which should be a generator
function).
* @param {Array} effort - An array representing the effort input for the
simulation.  Includes expected (distribution mode) and actual.
* @param {Array} revenue - An array representing the revenue input for the
simulation.  Includes expected (distribution mode) and actual.
*
* This function performs the following steps:
*
* 1. Validates that the `effort` and `revenue` arrays are of the same length. If
not, it throws an error.
* 2. Looks up the agent function (a generator) in the global context by the given
`agentName`.
*    - If the `agentName` does not exist or is not a generator function, it throws
an error.
* 3. Initializes the generator function (the agent) with the remaining arguments.
```

```
* 4. Iterates over the agent's yielded values and stores them in the `rtn` array.
It also checks if all values are `null` to determine if the agent generated only
null values.
* 5. Returns:
*    - `null` if all generated values were `null`.
*    - Otherwise, it sums the generated values and returns the total.
*
* Error Handling:
* - Ensures that the effort and revenue arrays have the same length.
* - Throws an error if the agent is not found or if it's not a valid generator
function.
*
* @returns {number|null} The sum of the values generated by the agent, or `null` if
all the values were `null`.
*
* Comment generated by ChatGPT. Code written by author.
*/
function runSim(agentName, effort, revenue){
 var args = [...arguments].slice(1,arguments.length);

 if(effort.length != revenue.length) {
   throw "Cost and Revenue arrays must be the same length";
 }

 // look up the agent (generator function) in the global namespace
 try {
   var agent = globalThis[agentName];
   if(!(agent instanceof (function*(){}).constructor)) {
     throw "Error";
   }

 } catch {
   throw "Simulation agent '"+agentName+"'"+" is not defined.";
 }

 // initialise the generator function (the agent)
 agent = agent(...args);

 // keep calling the agent until it completes and return the resulting array
 var rtn = [];
 var allNull = true;
 for (const value of agent) {
   rtn.push(value);
   allNull = allNull && value === null;
 }

 return allNull ? null :
   rtn.reduce((tot, x) => tot + x, 0);
}
```

Appendix F – Simulation Agents

Licence

```
/**
* R&D Agent Simulation (this appendix) © 2024 by Owen Woollaston is
licensed under Creative Commons Attribution-NonCommercial 4.0
International. To view a copy of this license, visit
https://creativecommons.org/licenses/by-nc/4.0/
*/
```

Code – ROI Agent

```
/**
* Generator function simulating an agent that makes project investment decisions
based on ROI thresholds and borrowing limits.
*
* @param {Array} effort - Array of effort data, where each entry is a tuple
[expected cost, actual cost].
* @param {Array} revenue - Array of revenue data, where each entry is a tuple
[expected revenue, actual revenue].
* @param {number} [roiThres=-1] - The minimum acceptable ROI threshold. Default is
-1 (allows any project with positive revenue).
* @param {number} [borrowingLimit=1e9] - The initial borrowing limit or bank
balance the agent starts with. Default is 1 billion.
*
* This generator function models an agent that processes a series of projects and
makes decisions about whether to commit to each project based on expected costs and
revenues.
*
* For each project (based on the provided `effort` and `revenue` arrays):
* 1. The agent calculates whether to commit to the project using the following
conditions:
*    - The expected cost is less than the current bank balance.
*    - The expected revenue exceeds the expected cost plus the ROI threshold.
*
* 2. If the agent does not commit to the project, it yields `null` to indicate
skipping the project.
*
* 3. If the agent commits but the actual cost exceeds the bank balance, it ends the
simulation early and returns the negative bank balance to indicate failure (running
out of funds).
*
* 4. If the project succeeds (i.e., the agent commits and has enough funds), the
profit is calculated as the difference between actual revenue and actual cost.
*    - The profit is added to the bank balance.
*    - The function yields the profit for that project.
*
* The function continues processing each project until all are completed or until
the agent runs out of money.
*
```

```
* @yields {number|null} - Yields the profit for successful projects or `null` for
skipped projects.
* @returns {number} - Returns the negative bank balance if the agent runs out of
money.
*
* Comment generated by ChatGPT. Code written by author.
*/
function* roiAgentWithBorrowLimit(effort, revenue, roiThres=-1, borrowingLimit=1e9)
{
 var bankBalance = borrowingLimit;

 for(let i=0; i<effort.length; i++) {
   var [expCost, actCost, expRevenue, actRevenue] = [...effort[i], ...revenue[i]];

   // decide whether to commit to this project or not
   var commitToProject =
     expCost < bankBalance &&         // we THINK there is enough cash in the bank,
and
     expRevenue > expCost * (1+roiThres);    // we THINK the ROI looks good

   if(!commitToProject) {
     yield null; // skip this project
   } else if(actCost > bankBalance) {
     return -bankBalance;  // we ran out of money before completing the project
   } else {
     // project was either a success or it didn't end the game.
     let profit = actRevenue - actCost;
     bankBalance += profit;
     yield profit;
   }
 }
}
```

Code – Risk-Aware Agent

```
/**
* Generator function simulating an agent that makes project investment decisions
based on ROI thresholds, borrowing limits, and risk thresholds.
*
* @param {Array} effort - Array of effort data, where each entry is a tuple
[expected cost, actual cost, risk].
* @param {Array} revenue - Array of revenue data, where each entry is a tuple
[expected revenue, actual revenue].
* @param {number} [roiThres=-1] - The minimum acceptable ROI threshold. Default is
-1 (accepts any project with positive revenue).
* @param {number} [borrowingLimit=1e9] - The initial borrowing limit or starting
bank balance. Default is 1 billion.
* @param {number} [riskThresh=1e9] - The maximum acceptable risk level for a
project. Default is 1 billion.
*
* This generator function models an agent that evaluates and decides whether to
commit to a series of projects based on expected costs, expected ROI, and risk
level.
```

```
*
* For each project (based on the provided `effort` and `revenue` arrays):
* 1. The agent evaluates whether to commit to the project using the following
conditions:
*    - The expected cost is less than the current bank balance.
*    - The expected revenue exceeds the expected cost plus the ROI threshold.
*    - The project's risk level is below the defined risk threshold.
* 2. If the agent decides not to commit to the project, it yields `null` to
indicate the project is skipped.
* 3. If the agent commits but the actual cost exceeds the available bank balance,
it returns the negative balance, indicating the agent has run out of money, thus
ending the simulation.
* 4. If the agent commits and the project succeeds, the profit is calculated as the
difference between the actual revenue and the actual cost. The profit is added to
the bank balance, and the profit is yielded.
* The function continues evaluating and committing to projects until all are
processed or until the agent runs out of money.
*
* @yields {number|null} - Yields the profit for successful projects or `null` for
skipped projects.
* @returns {number} - Returns the negative bank balance if the agent runs out of
money.
*
* Comment generated by ChatGPT. Code written by author.
*/
function* riskAwareAgentWithBorrowLimit(effort, revenue, roiThres=-1,
borrowingLimit=1e9, riskThresh=1e9) {
 var bankBalance = borrowingLimit;

 for(let i=0; i<effort.length; i++) {
   var [expCost, actCost, risk, expRevenue, actRevenue] = [...effort[i],
...revenue[i]];

   // decide whether to commit to this project or not
   var commitToProject =
     expCost < bankBalance &&           // we THINK there is enough cash in the
bank, and
     expRevenue > expCost * (1+roiThres) &&    // we THINK the roi looks good
     risk < riskThresh;                        // risk level looks good

   if(!commitToProject) {
     yield null; // skip this project
   } else if(actCost > bankBalance) {
     return -bankBalance;  // we ran out of money before completing the project
   } else {
     // project was either a success or it didn't end the game.
     let profit = actRevenue - actCost;
     bankBalance += profit;
     yield profit;
   }
 }
}
```

Code – Gated Agent

```
/**
* Risk-Aware Agent with Borrow Limit using stage-gate Process
*
* This generator function simulates decision-making for investments in projects
based on expected and actual effort, revenue, and associated risks. It uses a
two-phase stage-gate model for feasibility and development.
*
* @param {Array} effort - Array of effort values for projects, where each entry is
[expected cost, actual cost, risk].
* @param {Array} revenue - Array of revenue values for projects, where each entry
is [expected revenue, actual revenue].
* @param {number} [roiThres=-1] - Minimum ROI threshold for investment.
* @param {number} [borrowingLimit=1e9] - Maximum borrowing limit (initial bank
balance).
* @param {number} [riskThresh=1e9] - Maximum acceptable risk threshold for
projects.
* @param {number} [investment=1] - Proportion of expected cost to be invested in
Phase 1 (between 0.05 and 1).
* @param {number} [costImprovement=0.5] - Factor by which actual cost influences
updated expected cost.
*
* @yields {number} - Profit or loss from each project. If bank balance is
insufficient, returns a negative value indicating the shortfall.
*/
function* riskAwareAgentWithBorrowLimitUsingStagegate(effort, revenue, roiThres =
-1, borrowingLimit = 1e9, riskThresh = 1e9, investment = 1, costImprovement = 0.5)
{
   var bankBalance = borrowingLimit;

   // Ensure Phase 1 investment is within the valid range (5-100% of costs)
   let phase1Investment = Math.max(0.05, Math.min(1, investment));

   for (let i = 0; i < effort.length; i++) {
       let [expCost, actCost, risk, expRevenue, actRevenue] = [...effort[i],
...revenue[i]];

       /*****************************************************
        * PHASE 1 - FEASIBILITY
        *****************************************************/
       // Allocate Phase 1 costs as a proportion of expected costs
       let phase1Cost = phase1Investment * expCost;

       // Decide whether to commit to Phase 1
       var commitToPhase1 =
           expCost < bankBalance && // Sufficient funds in bank (expected)
           expRevenue > expCost * (1 + roiThres) && // ROI meets threshold
           risk < riskThresh; // Risk is acceptable

       if (!commitToPhase1) {
           // Skip this project completely
           continue;
```

```
        }

        /*****************************************************
         * PHASE 2 - DEVELOPMENT
         *****************************************************/
        // Re-estimate project cost with updated information
        let newExpCost = actCost * costImprovement + expCost * (1 -
costImprovement);

        // Review commitment to project based on updated estimates
        let commitToPhase2 =
            newExpCost < bankBalance && // Sufficient funds in bank (updated
estimate)
            expRevenue > newExpCost * (1 + roiThres); // Updated ROI meets threshold

        if (!commitToPhase2) {
            // Abandon project, deduct Phase 1 cost
            bankBalance -= phase1Cost;
            yield -phase1Cost;
            continue;
        } else if (actCost > bankBalance) {
            // Insufficient funds to complete project; exit simulation
            return -actCost;
        } else {
            // Project completed successfully, calculate profit
            let profit = actRevenue - actCost;
            bankBalance += profit;
            yield profit;
        }
    }
}
```

Appendix G – Evaluation of AI-Assisted Literature Search

The literature review method defines a systematic approach in which specific AI prompts were used to generate a numeric relevance rating to the research question. This is a potentially novel and unproven method of performing a literature search, so it is pertinent to demonstrate its performance. Specifically, the use of GPT-4o to score the relevance of an article (title and abstract) against a research question is investigated. This section aims to add rigour to the literature review by statistically testing the AI-generated results for (1) accuracy in terms of inter-rater reliability with this author and (2) repeatability across multiple ratings of the same article, that is, AI self-rater agreement.

Method

Following a similar method to that used to review AI-human inter-rater reliability by Landschaft et al. (2024) and the guidance from "Handbook of Inter-rater Reliability" (Gwet, 2014), a new search was run using the parameters in Table G1 and the method described in the literature review. The search is similar to that used for the primary thesis but generated a smaller dataset (202 articles).

Table G1

Test Search Parameters

Search Journals	As per Appendix A
Search Keywords	"research and development" AND "decision making" AND "product management"
Earliest Year	≥ 2010
Citation Percentile	≥ 95% (for articles older than 2010)
AI Review Keywords	decision-making, R&D, innovation management, product management, new product development, high-tech, escalation of commitment
AI Research Question	How should firms evaluate whether to continue or abandon projects in high-tech R&D?

Next, each article was assigned an AI-computed relevance (AI_R) score as part of the automated process. The data was downloaded to a spreadsheet, and a new column named "Human" was added to record the author's article rating. The *AI_R* column was hidden so that the AI-ratings could not bias this author. Then all 202 articles were rated manually using the Likert scale in Table G2 below. Articles rated zero (irrelevant) or one (adjacent subject) were deemed excluded, and articles rated two (rater was uncertain) or three (relevant) were deemed included. The time it took each rater to rate all 202 articles was noted, and all articles were rated in one sitting.

Table G2

Likert Scale for Human Rater

Ratings	Description	Determination
0 - irrelevant	Not relevant	Excluded
1 - adjacent	Of interest and related, but not relevant	Excluded
2 - uncertain	Possibly relevant - needs further reading	Included
3 - relevant	Clearly relevant	Included

Once the AI and human ratings were captured side-by-side in the spreadsheet, it was possible to determine inter-rater agreement. The *AI_R* column was unhidden. Then the distribution of the results was plotted to visualise the inter-rater agreement, and the count of the rejected and selected articles was plotted for values of *AI_R* exceeding a given threshold. Finally, Cohen's Kappa (a statistical measure of inter-rater agreement) was calculated for each integer value of *AI_R* (Cohen, 1960; Gwet, 2014).

To determine AI repeatability, three articles were chosen from the dataset for each value of *AI_R* as determined by the automated process to determine the repeatability. *AI_R* values from one to eight were observed – there were no nines or tens, so in all, 24 articles were selected. Then, for each chosen article, AI relevance (*AI_R*) was generated thirty times for a total of 720 samples. The results were visualised using a box and whisker plot, and Cronbach's Alpha (Cronbach, 1951; Gwet, 2014) was calculated to measure the results' internal consistency (repeatability).

Results and Discussion

It took this author 87 minutes to manually rate 202 articles – 26 seconds per article, most of which were screened by title alone and some by abstract if more information was needed. For the AI, it took 5 minutes to set up the search and 52 seconds to process all the titles and abstracts and calculate the AI relevance – about four articles per second and 100 times faster than this author.

Figure G1 below shows a histogram of the human ratings (left) and the AI-generated relevance (right). For these articles, GPT-4o returned *AI_R* scores ranging from one to eight. At first glance, there was some similarity in the shape – both approaches rated most articles as irrelevant.

Figure G1

Human Rating and AI-Relevance (AI_R)

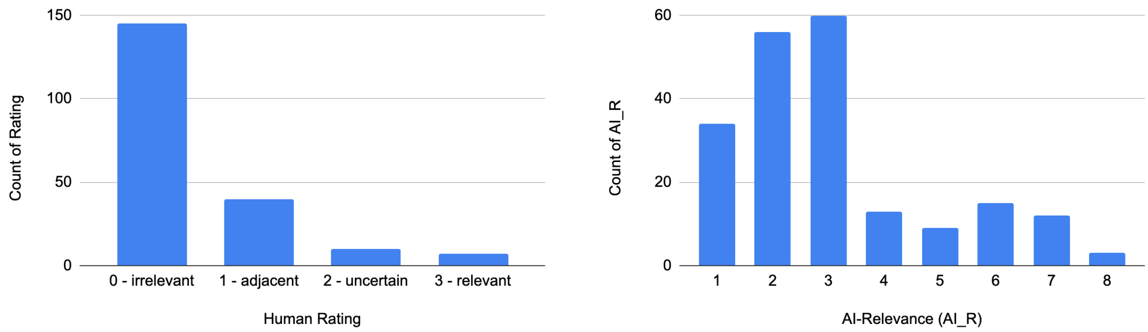


Figure G2 shows the number of articles AI selects and rejects for *AI_R* over a given threshold. The human rater selected 17 and rejected 185 articles. Low thresholds result in selecting too many irrelevant articles, and high thresholds result in rejecting too many relevant articles. Thresholds of four, five, and six indicate good agreement between human and AI raters.

Figure G2

AI Selectivity Versus Threshold

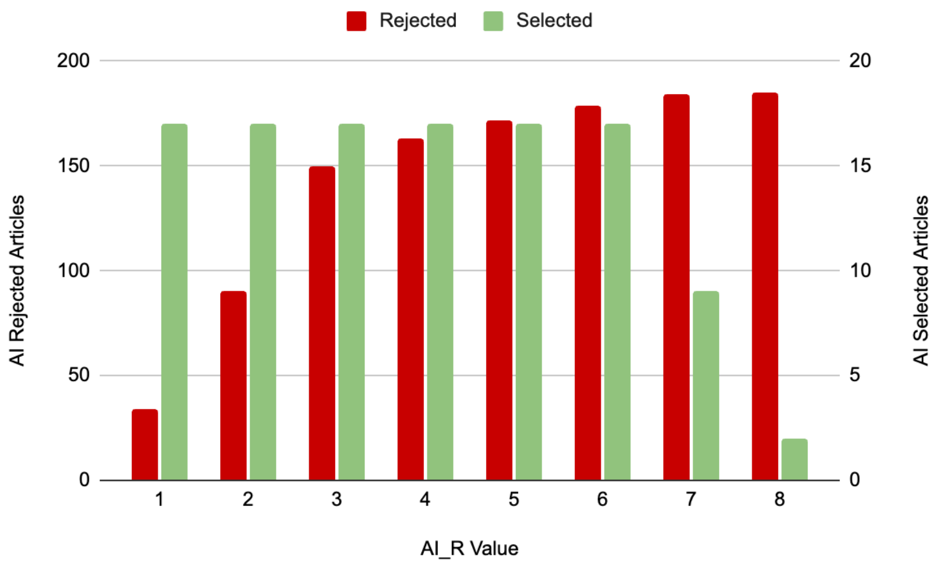


Table G3

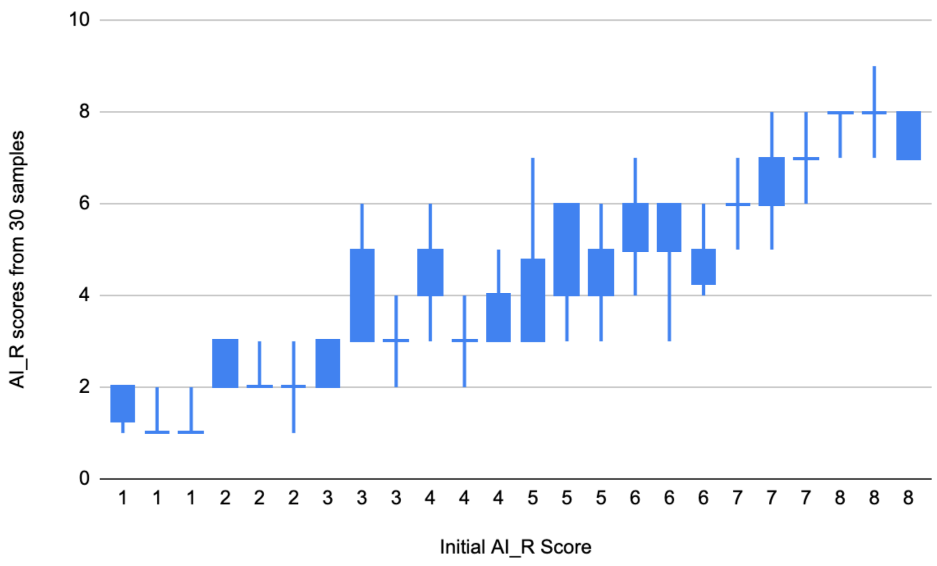
Inter-Rater Agreement

AI_R ≥	Cohen's Kappa	Agreement
1	0.128	Poor
2	0.158	Poor
3	0.241	Fair
4	0.479	Moderate
5	0.598	Moderate
6	0.717	Substantial
7	0.555	Moderate
8	0.196	Poor

By generating confusion matrices[7] for each *AI_R* threshold and calculating Cohen's Kappa, the optimal *AI_R* threshold can be identified. The results are shown in Table G3. Substantial agreement is achieved when selecting *AI_R* ≥ 6. It is important to note that *AI_R* ≥ 6 is the best threshold for this set of articles, and this author's experience with other searches indicates that *AI_R* ≥ 6 is a good default. However, other research questions may warrant a different threshold. In the case of the primary thesis, for example, *AI_R* ≥ 7 produced the optimal balance of workload and accuracy.

Figure G3

Box and Whisker Showing Spread of GPT Relevance Scores



[7] See the literature review approach for explanation of confusion matrices and examples.

Figure G3 shows box and whisker plots (vertical axis) for each article and *AI_R* score (horizontal axis). There are three articles for each initial *AI_R* score, so 24 articles are represented. The vertical axis represents the spread of the additional 30 *AI_R* values generated. Visually Figure G3 shows a strong relationship between the different AI_R samples, but with some variability. Further, when the entire dataset is used to calculate Cronbach's Alpha (a measure of internal repeatability), a value of 0.997 was determined, which is considered excellent. This result indicates that while intra-article ratings may vary somewhat, over a large set of articles the repeatability is excellent.

Variability might be a serious concern if the AI-assisted approach were to run unsupervised. However, as recommended by Burger et al. (2023), Assaf Landschaft et al. (2024), and Whelan et al. (2021), the process always ran under this author's supervision. In that respect, AI-assisted meant just that: assistance, not replacement.

Conclusion

In this appendix, the AI-assisted search method was evaluated for its accuracy and repeatability. The results show that using AI to implement a "better search" offers significant time savings and the ability to scale up the search size. If the correct threshold is chosen, the approach can be highly selective, reducing the number of irrelevant articles researchers must process manually without omitting relevant articles. Repeatability was excellent, as was the agreement between this author and AI. Further, Assaf Landschaft et al. (2024) used a similar approach and found good agreement between GPT-4 article rankings and independent human raters.

The ability to "cast the net wide" offered by this method has implications for the standard literature search process. The process of refining search parameters in searches (most commonly title, keyword, abstract, publication date) and specifying boolean combinations, therefore, is primarily aimed at reducing the researcher's effort in locating relevant articles. This method circumvents the need for early filtering and vastly increases the number of articles that can be initially screened for relevance. This author has run one search of 16000 articles, and many of 1000-2000 articles. Further, the standard guidance to restrict searches to recent years may be obsolete. Relevance is relevance. Publication date, keywords, and other parametric, boolean, exact, and wildcard searches are all weak proxies in comparison.

There are several opportunities for further research and improvements to this method:

1. **Multiple human raters** in a larger trial would add to the robustness of the statistical testing done here.

2. **Automated snowballing** is possible by inspecting article references, building a directed graph of references between articles and using the graph to identify key documents that should also be downloaded and rated.
3. **Further testing** using other research questions and topics could generate different results (accuracy or repeatability) depending on the language.

Beyond formal literature searches, the method provides a valuable learning and rapid-enquiry tool. When Google searches and ChatGPT sessions seem inadequate or need validation, the author’s favourite technique is to define and run an AI-assisted search and then review the top 10 most relevant articles. Further automation can be added by feeding the top 10 articles into ChatGPT and asking it to produce summary notes or digest specific content.

This field is rapidly evolving. Jin et al. (2024) and Souifi et al. (2024) review currently available tools to support this. As practitioners of literature searches and students of this emerging field, researchers would be well advised to get abreast of the available tools and the field’s future direction.

Appendix H – Adoption and Novelty of AI-Assisted Search

I am convinced that artificial intelligence will not replace humans. It will, however, replace those who refuse to work with it.

– Garry Kasparov

The literature review method defines a systematic approach in which specific AI prompts were used to generate a numeric relevance rating to the research question. By the time the primary thesis was drafted (October 2024), other researchers were exploring adopting AI into research methods. To understand how AI was being adopted to facilitate research, an AI-assisted search was used to produce a brief literature review using the parameters in Table H1. The search was configured to rank articles according to the question, “How widely adopted are AI-assisted literature search methods among academic researchers?”

Table H1

Search Parameters For AI-Assisted Search Adoption

Scopus Keywords	(literature search OR systematic review OR literature review) AND (AI-assisted OR AI-based OR AI-enhanced)
Earliest Year	2020
Journals	Any
Citescore	Any
AI Review Keywords	ai-assisted
AI Research Question	How widely adopted are AI-assisted literature search methods among academic researchers?
AI Relevance Cutoff	$AI\_R \geq 6.0$

The search results are summarised in Table H2 below. In total, 1534 articles were screened by AI, and 24 articles were identified. The term “machine learning” was omitted from the search in this review. There were two reasons for this. The first was that adding “machine learning” increased the number of Scopus search results from 1534 to 22645, a challenging number even for this AI-assisted method. This large result pool underscores the escalating problem facing literature searches – the volume of material is rapidly expanding beyond human capacity. The second reason was that “machine learning” refers to customised AI implementations, whereas an off-the-shelf AI service was used here. Another way to look at it is that AI – using large language models (LLMs) like ChatGPT – is built on machine learning.   The impact of this omission was to introduce the risk that similar methods are being implemented using machine learning, and Wagner et al. (2022) and Whelan et al. (2021) mention such methods. Typically, however, they use older approaches that predate the general availability of ChatGPT and require a higher understanding of programming. This method – using OpenAI's GPT-4o API – was comparatively simple to implement.

Table H2

Relevant Publications on AI-Assisted Reviews

Focus	Topic	Author/s
Method	Review of AI methods, How to use AI methods	Bolaños et al. (2024), Tóth & Oldal (2022), Wagner et al. (2022), Zala et al. (2024)
Ethics	Ethics for AI methods, Guidelines for AI methods	Burger et al. (2023), Lee et al. (2023), Mozelius & Humble (2024)
Performance	Evaluation of AI methods, Accuracy of AI-generated content, Impact on publishing ease and volume, Evaluation of AI written papers, Inaccuracies with AI-generated reviews, Performance of AI methods	Kacena et al. (2024), Margetts et al. (2024), Nazzal et al. (2024), Qureshi et al. (2023), Schmidt et al. (2022), Švab et al. (2023), Tomczyk et al. (2024), Yao et al. (2024)
Tools	Evaluation of AI research tools, Overview of AI search tools	Jin et al. (2024), Souifi et al. (2024)
Application	An AI-assisted method, Application of AI as reviewer, Specific application of AI to search, Visual search and exploration using AI, Using AI for different types of review	Abd-Alrazaq et al. (2021), Ferrati et al. (2024), Landschaft et al. (2024), Sina et al. (2024), Vihari & Kaur (2024), Whelan et al. (2021)
Trends	Trends in AI writing	Roy et al. (2024)

From these results, it can be seen that a few researchers have published using AI-assisted methods in medicine, biogenetics and entrepreneurship (Abd-Alrazaq et al., 2021; Ferrati et al., 2024; Landschaft et al., 2024; Sina et al., 2024; Vihari & Kaur, 2024; Whelan et al., 2021), and according to Bolaños et al. (2024), the prevalence of researchers integrating Large Language Model (LLM) AI, such as ChatGPT has risen dramatically since 2022.

Various methods are reported, spanning all phases of the systematic literature review process: planning, search, selection, data extraction and synthesis, quality assessment and reporting. The general theme is that ethical considerations become more complex with each additional AI phase (Burger et al., 2023; Lee et al., 2023; Mozelius & Humble, 2024). At one extreme, AI can assist with planning by suggesting structure and refinements to research questions and keywords. This raises no ethical concerns because all the work will be original and done by the author, and the role of AI is advisory – like having a digital academic supervisor. On the other hand, researchers are testing the performance of AI in writing entire articles. These articles typically lack accuracy, particularly regarding specific facts and references (Margetts et al., 2024; Kacena et al., 2024).   In this context, the work is ethical because AI is the subject of study. However, using AI this way would be highly unethical for most researchers because the human contribution would be minimal, and the work would not be original or authentic. Nonetheless, the efficiency improvement is indisputable, with researchers reporting a 20-88% reduction in effort (Yao et al., 2024). In May 2024, while developing this method, Elsevier released Scopus AI, aimed at improving the efficiency and efficacy of searches:

> "Scopus AI is an intuitive and intelligent search tool powered by generative AI (GenAI) that enhances your understanding and enriches your insights with unprecedented speed and clarity." (Elsevier, 2024)

When this thesis began in early 2024, the AI-assisted literature search method presented here was highly novel, with no prior publications detailing a similar approach. Since then, the rapid integration of AI into research workflows has led to increasing adoption, with a few researchers now publishing method descriptions (e.g., Joos et al., 2024).

While this has somewhat eroded the exclusivity of the approach, this method remains novel in several key aspects:

1. The structured AI prompt design that generates numeric ranking scores for relevance.
2. The systematic reliability testing (Appendix G), which validates AI consistency.
3. The practical integration of AI into a replicable screening workflow.

Thus, while AI-assisted literature search methods are gaining traction, the specific contributions of this approach – including its quantitative ranking mechanism and validation process – continue to represent a meaningful contribution to research methodology.