MANAGEMENT
2026, Vol. 30, No. 1
www.management-poland.com
DOI: 10.58691/man/195476

ENGIN YURDASEVER
CANAN YILDIRAN

GenAI-Supported strategy development in VUCA scenarios and Multi-GenAI evaluation

Abstract

Research background and purpose: Today's VUCA environment-characterized by increasing volatility, uncertainty, complexity, and ambiguity-is making the limits of both cognitive and analytical capabilities more apparent in organizations' strategy development processes. The aim of this study is to examine how the strategy development and strategic reasoning capabilities of generative artificial intelligence (GenAI) tools differ across individual VUCA components.

Design/methodology/approach: The research employs an exploratory design that combines a qualitative scenario method with multi-GenAI-based strategy development and an Artificial Intelligence-Based Evaluation (AI-as-Evaluator) approach. In line with the research objective, four separate scenarios representing each component of VUCA were generated by the generative AI tool called ScholarGPT, and within the context of these scenarios, strategy proposals were developed by the generative AI tools ChatGPT, Gemini, DeepSeek, and Claude. The strategies developed were evaluated by multiple generative AI tools, within the framework of the principles of blindness and non–self-scoring, in terms of the criteria of innovativeness and creativity, feasibility, agility and adaptability, risk level, and market alignment.

Findings: The findings reveal that although generative AI tools exhibit high performance in "agility and adaptability" across all scenarios, the "complexity" component leads to a systematic decline in the "feasibility" scores of the strategies and a marked disruption in evaluator consistency.

Value added and limitations: The study expands the VUCA literature from an artificial intelligence perspective and offers a conceptual and methodological framework for research on generative AI-supported strategy development.

Keywords: *VUCA, generative AI (GenAI), AI-supported strategy, Multi-AI evaluation, scenario planning*

JEL Classification: M10, M15, D81

Received: 2026-03-2; **Revised:** 2026-04-20; **Accepted:** 2026-04-27

Engin Yurdasever
Social Sciences Vocational School, Ordu University, Turkey
ORCID: 0000-0002-3853-2032

Canan Yıldıran
Faculty of Business, Karabük University, Turkey; email: cananyildiran@karabuk.edu.tr
ORCID: 0000-0001-8245-197X

1. Introduction

In today's business world, the high level of uncertainty that has emerged with the increase in global interconnectedness and the pace of change is widely defined in the literature by the acronym VUCA (volatility, uncertainty, complexity, and ambiguity) (Bennett & Lemoine, 2014b). Under current conditions, where change has become continuous, VUCA (Volatility, Uncertainty, Complexity, Ambiguity) is no longer seen as a temporary situation but is regarded as the "new normal" of the business world (Sullivan, 2012). In this new normal, the prevailing extreme uncertainty and the inability to establish clear cause–effect relationships render traditional planning methods inadequate and make it significantly more difficult to make rational strategic decisions and to formulate strategy (Bennett & Lemoine, 2014a; Thorén & Vendel, 2019).

It is known that each of the components that constitute VUCA-volatility, uncertainty, complexity, and ambiguity-contains its own distinctive strategic challenges that threaten organizational performance (Rzepczynski, 2025; Shet, 2024). However, it is observed that in the literature, in-depth studies that analyse these components separately and in a disaggregated manner are still quite limited in number (Taskan et al., 2022).

The extreme uncertainty and ambiguity of cause-effect relationships inherent in VUCA environments severely constrain decision makers' ability to act in a fully rational manner, and this situation is closely associated with the theory of bounded rationality, which posits that decisions are made under cognitive, temporal, and informational limitations (Simon, 1991; Rzepczynski, 2025). On the other hand, the increase in environmental dynamism and volatility raises the amount of information organizations need to make sound decisions, and this is linked to information processing theory, which explains firms' decision-making processes through their information processing capacity (Galbraith, 1977).

Today, artificial intelligence (AI) technologies have rapidly expanded their areas of application by becoming an integral part of strategic planning and decision support systems, thanks to their capabilities for collecting and analysing data (Perifanis & Kitsios, 2023). Especially with the development of generative AI systems, these tools are no longer limited to merely analysing data; they are transforming into active "decision partners" that can reason by deriving meaning from complex data and offer strategic recommendations (Kesgin, 2025; Takemoto, 2024; Almutairi & Almatrodi, 2025). While these developments provide a technological opportunity to support the limited capacity of the human mind, they also create a highly natural and powerful common ground with the bounded rationality and information processing framework that forms the theoretical basis of this study (Pittenger et al., 2023). However, knowledge regarding how each of these advanced tools affects the distinct VUCA challenges remains quite limited (Kitsios & Kamariotou, 2021).

This study aims to evaluate the strategy recommendations generated by different generative AI tools, within the context of business scenarios constructed based on each VUCA component, in terms of predetermined performance criteria; to reveal the similarities and the dimension-based systematic divergences of these recommendations; and to examine the consistency of evaluations across generative AI tools. In this respect, the study seeks to identify in what ways the strategic reasoning and strategy development capacities of generative AI tools differ under the various components of VUCA, and it aims to fill an important methodological and conceptual gap in the literature on AI-supported strategy development.

In summary, the study theoretically re-examines the VUCA literature from an AI perspective and expands information processing theory and the bounded rationality approach in the context of developing strategies supported by AI. From a methodological standpoint, by employing the AI-as-Evaluator approach, which enables the analysis of strategy recommendations through a blind and multiple-evaluation logic, it offers an original assessment framework for AI-based comparative strategy research.

2. Theorical background

2.1. Volatility, Uncertainty, Complexity, Ambiguity

The acronym VUCA was first used in the late 1990s by the U.S. Army War College to describe the multipolar and fragmented world order that emerged after the Cold War (Kinsinger & Walch, 2012; Yarger, 2006). This military-origin acronym was rapidly adopted in the management and strategy literature, particularly with the 2008 financial crisis and the subsequent waves of global uncertainty and has become a central framework for making sense of the dynamic nature of modern markets (Baran & Woznyj, 2021; Bennett & Lemoine, 2014a). However, among the main criticisms in the literature are that the concept is sometimes used as an overly broad metaphor and that the boundaries between its components are blurred, which in turn brings about difficulties in measurement and application (Mack & Khare, 2016; Taskan et al., 2022). To overcome the complexity created by this conceptual breadth, an increasingly advocated approach in the literature is to analyse each component separately rather than treating VUCA as a holistic construct, thereby providing a clearer framework for strategy development processes (Bennett & Lemoine, 2014b; Winkler et al., 2025).

In the literature, volatility is defined as an increase in the frequency of environmental changes and the unstable nature of these changes (Sullivan, 2012), whereas uncertainty represents situations in which, even if the underlying cause-effect relationships are known, the capacity to make meaningful predictions about the trajectory of future events and the outcomes they will generate has been lost (Bennett & Lemoine, 2014b; Horney et al., 2010). Complexity refers to the presence of a large number of interrelated parts and variables that constitute a system or situation (Bennett & Lemoine, 2014b; Thorén & Vendel, 2019), while ambiguity denotes a domain of uncertainty for which there is no historical precedent and in which decision makers cannot fully define what they are facing (Bennett & Lemoine, 2014b; Rzepczynski, 2025).

2.2. Theoretical framework

This study brings together Information Processing Theory and the Bounded Rationality approach within a common framework to make sense of strategy formulation and evaluation processes under the rapidly changing conditions induced by VUCA environments. When these two theoretical perspectives are considered jointly, the misalignment between the intense information requirements that arise as environmental uncertainty increases and the cognitive limitations of the decision makers who must process this information is viewed as the key factor determining the quality of strategic decisions (Biloslavo et al., 2025; Luo & Donaldson, 2013).

At the core of Information Processing Theory lies the assumption that as the complexity of the tasks an organization must perform or the level of environmental uncertainty increases, the amount of information required to cope with this situation will also rise (Galbraith, 1977). According to this theoretical expectation, in a VUCA environment, particularly as complexity and uncertainty increase, it is not sufficient for an effective strategy recommendation to merely present data. This data must be prioritized, meaningfully structured, and accompanied by an explicit articulation of the underlying strategic assumptions (Yang et al., 2025). Therefore, today, AI is expected to balance the growing need for information processing and thereby reduce the cognitive load on decision makers (Biloslavo et al., 2025).

Bounded Rationality Theory emphasizes that the human mind has a limited capacity to evaluate complex problems in all their dimensions, and that most decisions are made under conditions of time pressure, incomplete information, and high mental load (Simon, 1955). Therefore, according to the theory, instead of full rationality aimed at finding the optimal solution, decision makers tend to choose solutions that are "good enough" under the prevailing circumstances (Simon, 1997; Vargas-Hernández & Pérez Ortega, 2019). This makes bounded rationality more pronounced in VUCA environments.

2.3. AI-Supported strategy development

In organizations, the classical strategy development process has a structure that requires intensive information processing and consists of stages such as environmental analysis, the formulation of strategic objectives, and the generation and evaluation of alternatives (Csaszar et al., 2024). At this stage, AI is positioned as a supportive tool that enhances the organization's information processing capacity, rapidly scans alternatives, and extends the bounds of rationality (Shrestha et al., 2019; Liu et al., 2023). The use of generative AI in the strategy domain, unlike traditional decision support systems that focus on numerical data, has enabled it to assume active roles in strategy formulation and evaluation processes by virtue of its ability to process qualitative and unstructured textual data (Csaszar et al., 2024). AI-supported strategic planning holds strong potential for rapidly generating many scenarios to save time, providing objective analyses free from human emotions and biases, and producing original strategic options from complex data sets (Duan et al., 2019). Thanks to this potential, it rapidly identifies complex patterns in large datasets, provides decision-makers with data-driven proactive insights, and increases the speed of operational decision-making (Biloslavo et al., 2024). Generative models, on the other hand, have the capability to instantly generate many strategic scenarios and to accelerate planning cycles by analysing these options with a depth comparable to that of human experts (Csaszar et al., 2024). However, the success of these tools is constrained by factors such as organizational design, data quality, and implementation context. For example, the recommendations offered by AI systems based on standardized data sets may trigger the risk of "decision isomorphism" thereby weakening strategic distinctiveness, and the fact that these tools cannot fully grasp tacit factors such as leadership vision and corporate culture carries the potential to cause strategic blindness in the face of radical market changes (Benhür Aktürk, 2025; Bessa & Barbosa, 2025). Accordingly, there remains a significant gap in the literature regarding the specific effects of AI on corporate strategy (López-Solís et al., 2025).

2.4. Strategy evaluation criteria in VUCA contexts

Strategic evaluation is defined as the process of aligning an organization's internal capabilities with environmental demands to ensure the efficient allocation of corporate resources (Tavana & Banerjee, 1995). This process aims to make the most rational choice under uncertainty by systematically analysing the potential risks and returns associated with alternative strategies (Steiner, 1969). Since VUCA contexts make it difficult to make strategic decisions based solely on rationality, organizations are turning to multidimensional evaluation criteria that consider both the process and the outcomes, rather than relying on individual indicators, when assessing strategic performance (Wei, 2025). In this context, it is evident that certain evaluation criteria in the literature can make the specific strategic challenges posed by the VUCA components more measurable. One of these is the innovation and creativity criterion, which captures the capacity of a strategy to generate new assumptions and to question existing paradigms when the market logic has not yet become clear (Teece et al., 2016); the feasibility criterion, which concerns clarifying operational steps and ensuring that the strategy is implementable (Hastings, 1996; Rumelt, 1979); the agility and adaptability criterion, which reflects the ability to rapidly review strategies and flexibly reallocate resources (Doz & Kosonen, 2010); the risk level criterion, which evaluates the extent to which the strategy is resilient to potential external shocks (Wei, 2025); and the market alignment criterion, which assesses the alignment of the strategy with market conditions (Rumelt, 1979).

2.5. AI-as-Evaluator

The AI-as-Evaluator approach, in which Large Language Models (LLMs) assume an evaluative role, offers significant operational advantages over traditional expert evaluations, such as high scalability, standardization capacity, and cost-effectiveness (Jacob, 2025). Especially in complex scenarios, AI possesses the capability to conduct large-scale simulations that can uncover regulatory loopholes and strategic trends that manual testing cannot detect (de Mesentier Silva et al., 2017). However, this approach also entails risks such as randomness in evaluation processes, systematic biases, and "evaluation-driven convergence," defined as the tendency for evaluation outputs to become confined to a specific pattern. Therefore, as in this study, adopting the principle of blind evaluation (Gu et al., 2025) and presenting outputs to the tools in a structured format (Laskar et al., 2025) stand out as measures that help mitigate this risk.

The theoretical perspectives and methodological tools addressed up to this point complete the analytical ground upon which the study is built and provide a holistic perspective on how the strategic reasoning process in the VUCA context will be measured. To summarize, throughout the study VUCA is not treated as a single, but rather as a set of contexts in which each component creates different problem structures and information-processing requirements. The Information Processing and Bounded Rationality perspectives provide a conceptual foundation for explaining the cognitive and analytical limitations of AI tools in both the formulation and evaluation stages of strategies developed within these contexts. In line with this, the performance criteria and the AI-as-Evaluator approach operationalize these theoretical assumptions, making them measurable and enabling the analysis of how strategy recommendations differ across various VUCA scenarios and whether this differentiation forms systematic and consistent patterns across dimensions.

3. Methods

3.1. Research design

The study employs an exploratory research design that combines a qualitative scenario method with multi–generative AI-based strategy formulation and generative AI-based blind evaluation. The main objective of the study is to examine how the strategy development and strategic reasoning capabilities of generative AI tools differ among the individual VUCA components. In line with this main objective, the research has been structured around sub-objectives so that the findings can be analysed in a multidimensional and comparative manner:

1. To analyse how generative AI-based strategy recommendations for business scenarios constructed based on different VUCA components differ in terms of the evaluation criteria (innovation, feasibility, agility, risk, and market alignment).
2. To identify the systematic patterns observed in the strategy outputs developed for different VUCA scenarios based on the evaluation criteria.
3. Within the framework of the AI-as-Evaluator approach, examine the level of consistency in strategy evaluations among GenAI-based evaluation tools.

The research process consists of three sequential stages. The first stage involves the generation of scenarios based on the VUCA components; the second stage comprises the development of strategy recommendations for these scenarios using generative AI tools; and the third and final stage consists of the blind evaluation and scoring of the developed strategies through the AI-as-Evaluator approach. The overall flow of the research, the relationships between the stages, and the roles of the generative AI tools are presented in Figure 1.

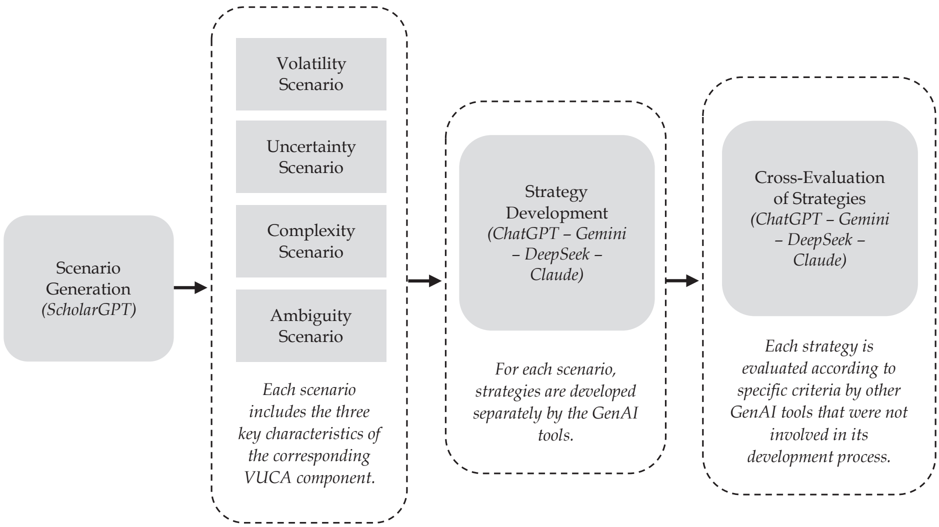

Figure 1. **Research flow and the role distribution of GenAI tools**

Source: own study

The tools used in the study were accessed via their web interfaces. The tools/versions employed are as follows: ScholarGPT v5.2 (scenario generation), ChatGPT v4.1, Gemini v3, DeepSeek v3.2, and Claude 4.5 Sonnet (strategy development and strategy evaluation). Access to the tools and the generation of outputs were carried out between 1–4 February 2026. In the implementation phase, the following core control mechanisms were adopted to ensure replicability and consistency: initiating a new and independent chat for each output, using a standardized prompt at each stage, and storing the outputs in their raw form.

3.2. Scenario method and scenario generation process

The scenario method is regarded as an approach that stimulates strategic thinking, particularly in decision environments characterized by uncertainty, complexity, and rapid environmental change (Wulf et al., 2012). Since VUCA also describes contexts in which the future is unpredictable and linear planning approaches are of limited use (Biloslavo et al., 2025), the scenario method provides a suitable methodological framework for concretizing strategic decision contexts by disentangling each component (Gürsoy & Şavk, 2024; Oral, 2024). In this study, a total of four scenarios were constructed, with one organizational scenario for each VUCA component. The function of the scenarios is to serve as "decision stimuli" to be used in the subsequent stage of strategy formulation. The scenarios were generated using ScholarGPT v5.2, one of the advanced generative AI tools. All scenarios were developed with a single, highly standardized scenario-generation prompt. The scenarios were designed as top-management-level (CEO/executive board) situations requiring strategic decisions, and for each scenario it was required that the three key features of the respective VUCA component most emphasized in the literature be included in a clear, concrete, and distinguishable manner. The authors have verified and approved whether the scenarios include the relevant features. The four scenarios developed as part of this study and their key features are presented in Table 1.

Table 1. **VUCA Scenarios Generated by ScholarGPT**

VUCA component	Scenarios	Key features
Volatility	The executive board of a major commercial bank is forced to rapidly restructure its balance sheet due to sudden, unpredictable fluctuations in foreign exchange rates and interest rates that shift direction within hours, as well as a liquidity notice issued overnight; significant changes in customer demand narrow the decision-making window and increase risks.	Speed and Dynamics of Change (Lawrence, 2013) Instability and Unpredictability (Sullivan, 2012) Volume and Magnitude of Change (Bader et al., 2019)

Uncertainty	A large hospital chain is facing uncertainty regarding reimbursement rates, conflicting demand projections, and cost fluctuations; management cannot fully assess the significance and strategic implications of these developments, and as decision-making timeframes shrink, liquidity and reputational risks are increasing.	Unpredictability and Lack of Foresight (Bennett & Lemoine, 2014a) Lack of Knowledge and Information Processing Requirements (Cernega et al., 2024) Uncertainty of Meaningful Ramifications and Significant Change (Fridgeirsson et al., 2021)
Complexity	A global energy company is struggling to manage its production portfolio due to the combined impact of price indices, infrastructure delays, and revisions to incentives; the CEO's decision-making window is limited due to conflicting demand, a heavy data flow, nonlinear relationships, and extreme complexity.	Multitude and Interconnectivity of Parts (Shet, 2024) Non-linear Dynamics and Emergence (Syamsir et al., 2025) Information Density and Constraints in Processing Capacity (Taskan et al., 2022)
Ambiguity	A mid-sized AI company faces rising demand, uncertain regulations, and sudden cancellations; although the data is rich, the cause-and-effect relationships and the rules governing the business model are unclear, and while management members offer differing interpretations, delayed decisions threaten investor confidence and operational sustainability.	Unclear Causal Relationships and "Unknown Unknowns" (Rzepczynski, 2025) Multiple and Conflicting Interpretations (Biloslavo et al., 2025) Lack of Knowledge of the "Rules of the Game" and Difficulty in Conceptualization (Yawson & Goryunova, 2025)

Source: own study

3.3. Strategy development process

In the strategy development stage, each scenario was presented separately to four different generative AI tools using the same standardized strategy-development prompt. In this study, both in the strategy development and strategy evaluation stages, four different generative AI tools (ChatGPT, Gemini, DeepSeek, Claude), representing different developer ecosystems and model families, were employed. The generative AI tools were asked to produce a single strategy recommendation that is short, clear, concrete, and comparable, and that fits the situation described in the scenario. Strategy development was carried out in a single-shot manner for each scenario-tool pair. Some example strategies developed by the generative AI tools within the scope of the study are as follows:

The strategies developed by generative AI tools within the scope of this study are shown in Table 2.

Table 2. **Strategies Developed by Generative AI Tools**

Type of strategy	Developer generative AI tool	Strategy
Volatility strategy	ChatGPT	Implementation of a dynamic risk framework and flexible decision-making mechanisms that can respond instantly to sudden market and regulatory shocks.
	Gemini	Abandoning fixed targets and adapting to market fluctuations through a rapid-intervention liquidity regime and an agile capital strategy.
	DeepSeek	Enhancing agility in liquidity and capital management through dynamic resource optimization and high-frequency market signals.
	Claude	Adapting to market shocks through flexible capital and liquidity buffers and automated real-time decisions.
Uncertainty strategy	ChatGPT	A strategy focused on flexible portfolios and financial resilience against uncertain regulations and conflicting demand signals.
	Gemini	Adapting to uncertain regulations through dynamic capacity and liquidity management, while remaining flexible and financially resilient.
	DeepSeek	Rapidly scaling resources with flexible focus, responding to uncertainty, and increasing operational agility.
	Claude	Adapting to uncertain regulations through flexible capacity and portfolio diversification, while preserving liquidity and strategic flexibility.
Complexity strategy	ChatGPT	Managing interdependent variables within a single framework, tolerating conflicting data, and balancing the short and long term.
	Gemini	Reducing data inconsistencies and adapting to complex market dynamics through a cross-functional decision-making model and flexible capacity management.
	DeepSeek	Managing conflicting data and interdependent variables through an integrated scenario platform, making decisions more forward-looking and aligned.
	Claude	Managing data conflicts and preserving strategic flexibility through an integrated decision infrastructure and risk-weighted capacity allocation.
Ambiguity strategy	ChatGPT	Using uncertainty as a strategic opportunity and achieving sustainable differentiation through adaptive infrastructure and risk-focused positioning.
	Gemini	Turning uncertainty into an opportunity through controlled agility and regulatory leadership, achieving alignment and flexibility through strategic differentiation.
	DeepSeek	Manage uncertainty, test growth opportunities, and mitigate risks through controlled agile innovation and a modular MVP (Minimum Viable Product) portfolio.
	Claude	Managing uncertainty and balancing flexibility with opportunities through regulatory compliance and controlled expansion into pilot markets.

Source: own study

The strategies presented in Table 2 have been generated by different artificial intelligence tools in response to the relevant VUCA scenarios and serve as actionable and comparable recommendations for decision-makers.

3.4. Evaluation of strategy performance: Evaluation criteria and scoring

The AI-as-Evaluator approach was adopted for the evaluation of the strategies developed (de Mesentier Silva et al., 2017; Gu et al., 2025). In this approach, generative AI tools are positioned as evaluators that score strategy recommendations based on predefined criteria. The evaluations conducted by the generative AI tools were carried out in line with the principle of blindness. The strategy texts were presented to the evaluating tools only together with the scenario context, and at no stage was information disclosed about which tool had generated the strategy. In addition, in accordance with the no self-scoring principle, no generative AI tool evaluated its own strategy (Li et al., 2024). The implementation of this principle and information on the roles of the tools in the study is summarized in Table 3.

Table 3. **Roles of generative AI tools**

Tools	Scenario generation	Strategy development	Strategy evaluation			
			ChatGPT	Gemini	DeepSeek	Claude
ScholarGPT	✓	✕	✕	✕	✕	✕
ChatGPT	✕	✓	✕	✓	✓	✓
Gemini	✕	✓	✓	✕	✓	✓
DeepSeek	✕	✓	✓	✓	✕	✓
Claude	✕	✓	✓	✓	✓	✕

Note: ✓: possible, ✕: not possible

Source: own study

As shown in Table 3, the scenarios were generated exclusively by ScholarGPT, while the strategies were developed by ChatGPT, Gemini, DeepSeek, and Claude, and in the evaluation stage none of the tools assessed the strategies it had produced itself. For use in the strategy evaluation phase, five fundamental evaluation criteria that are most emphasized in the literature for assessing strategy performance have been identified. These criteria are listed as (i) Creativity and innovation (Teece et al., 2016) (ii) Feasibility (Hastings, 1996; Rumelt, 1979) (iii) Agility and adaptability (Doz & Kosonen, 2010) (iv)

Level of risk (Wei, 2025) (v) Market alignment (Rumelt, 1979). In selecting the criteria, care was taken to synthesize the fundamental models in the strategic management literature with modern multi-criteria decision-making.

During the evaluation phase, numerical scoring between 1 and 10 was requested from the tools for each criterion. For each criterion, the arithmetic mean of the evaluation scores provided by the GenAI. tools was accepted as the final criterion score of the relevant strategy. In addition, standard deviation values were reported to describe the consistency across evaluations. Within this scope, the data collected from the generative AI tools were systematically analysed and are presented in the findings section.

4. Findings

In this section, the evaluation results for the strategy recommendations developed by generative AI tools for the scenarios representing different components of VUCA are presented.

4.1. Strategy evaluation findings by VUCA components and evaluation criteria

In Table 4, all evaluation scores for the strategy recommendations developed by the generative AI tools for the scenarios representing different components of VUCA are presented together, and these scores form the basis for the mean and standard deviation calculations used in the subsequent analyses.

Table 4. **Raw generative AI-based evaluation scores of strategies**

	Type of strategy																				
	Volatility					**Uncertainty**					**Complexity**					**Ambiguity**					
Strategy (GenAI Tool)	Evaluator	I	F	A	R	M	I	F	A	R	M	I	F	A	R	M	I	F	A	R	M
ChatGPT	Gemini	8	5	9	7	8	7	6	9	8	8	7	4	8	5	6	8	6	9	7	8
	DeepSeek	7	5	9	6	8	7	6	8	8	7	7	4	6	5	5	8	4	9	6	7
	Claude	7	5	8	6	7	7	4	8	6	7	7	5	6	6	7	7	5	8	6	7
Gemini	ChatGPT	6	7	8	6	7	6	7	8	7	6	6	7	8	6	7	7	6	8	7	6
	DeepSeek	7	6	9	5	7	8	5	9	7	6	3	2	6	4	5	7	6	9	8	7
	Claude	7	6	8	7	5	6	5	7	7	5	6	4	7	5	6	7	6	8	7	6

DeepSeek	ChatGPT	9	6	10	5	8	8	6	9	7	8	8	6	9	5	7	8	6	9	7	8
	Gemini	8	5	9	6	8	8	5	9	6	7	7	5	8	6	8	7	5	8	6	7
	Claude	8	4	9	3	6	7	4	8	5	6	7	4	8	5	6	7	5	8	6	7
Claude	ChatGPT	8	6	9	7	8	7	6	8	6	7	8	6	9	7	8	7	8	6	7	6
	Gemini	7	5	9	6	8	4	7	8	6	9	6	5	8	7	8	4	8	6	7	5
	DeepSeek	8	6	9	7	8	6	8	9	7	8	7	4	8	5	6	5	7	6	7	6

I: innovation and creativity, F: feasibility, A: agility and adaptability, R: risk level, M: market alignment

Source: own study

In Table 4, the results of 240 separate evaluations conducted by the generative AI tools for the strategy recommendations developed within the VUCA scenarios are presented. The scores reported in the table indicate that the evaluations of different generative AI tools regarding the strategy recommendations are not uniform and display a certain degree of variation. This variation is observed both across different VUCA scenarios and across evaluation criteria, and it provides an initial overview that helps in interpreting the mean and standard deviation findings addressed in the subsequent analyses.

4.2. Findings on criterion-level mean scores of strategies

In Table 5, the mean scores that the strategy recommendations developed within the VUCA scenarios received from the generative AI tools based on the evaluation criteria are compared, and the overall patterns that emerge at the criterion level are described.

Table 5. **Mean evaluation scores of strategies**

	Volatility strategy						Uncertainty strategy					
	I	F	A	R	M	OA	I	F	A	R	M	OA
ChatGPT	7,33	5,00	8,67	6,33	7,67	7,00	7,00	5,33	8,33	7,33	7,33	7,07
Gemini	6,67	6,33	8,33	6,00	6,33	6,73	6,67	5,67	8,00	7,00	5,67	6,60
DeepSeek	8,33	5,00	9,33	4,67	7,33	6,93	7,67	5,00	8,67	6,00	7,00	6,87
Claude	7,67	5,67	9,00	6,67	8,00	7,40	5,67	7,00	8,33	6,33	8,00	7,07

	Complexity strategy						Ambiguity strategy					
	I	F	A	R	M	OA	I	F	A	R	M	OA
ChatGPT	7,00	4,33	6,67	5,33	6,00	5,87	7,67	5,00	8,67	6,33	7,33	7,00
Gemini	5,00	4,33	7,00	5,00	6,00	5,47	7,00	6,00	8,33	7,33	6,33	7,00
DeepSeek	7,33	5,00	8,33	5,33	7,00	6,60	7,33	5,33	8,33	6,33	7,33	6,93
Claude	7,00	5,00	8,33	6,33	7,33	6,80	5,33	7,67	6,00	7,00	5,67	6,33

I: innovation and creativity, F: feasibility, A: agility and adaptability, R: risk level, M: market alignment, OA: overall average

Source: own study

An examination of Table 5 shows that the strategy recommendations developed for the VUCA scenarios differ markedly in terms of the evaluation criteria. When the overall average (OA) values are considered, it is observed that performance scores are relatively higher for Volatility and Uncertainty strategies (6.60-7.40), whereas lower performance values emerge for all generative AI tools in Complexity strategies (5.47-6.80). In Ambiguity strategies, by contrast, the overall means appear to be at a balanced level, with only limited differences in performance levels and distributions across tools.

When the data presented in Table 4 are examined at the criterion level, it is observed that the differentiation among VUCA strategies is particularly concentrated in the "Feasibility (F)" criterion. In Complexity strategies, feasibility scores cluster within a low band across all generative AI tools (4.33-5.00), whereas in Volatility and Ambiguity strategies they display a wider and relatively higher distribution. Similarly, it is noteworthy that the "Risk Level (R)" scores are more compressed at lower values in Complexity strategies (5.00-6.33). Although the differentiation across strategies remains more limited for the other evaluation criteria, the findings show that VUCA components distinguish the strengths and weaknesses of strategy outputs at the criterion level.

When these findings are considered together, the most systematic pattern that emerges is that the Complexity component constitutes a challenging and constraining context, regardless of which generative AI tool is used. For the strategies developed based on this component, not only do the overall mean scores decline, but the Feasibility (F) scores also exhibit a downward clustering across all generative AI tools. A similar, though more limited, pattern is observed for the Risk Level (R) criterion. By contrast, the fact that the Agility and Adaptability (A) criterion displays relatively high and more stable performance in almost all strategies indicates that generative AI tools are more inclined to develop agility- and adaptation-oriented strategies across different contexts. Finally, while the overall performance averages of the generative AI tools remain relatively close to each other in Volatility and Uncertainty strategies, the fact that these differences become more pronounced in Complexity strategies reveals that conditions of high complexity make divergences in strategic reasoning capacities more visible.

4.3. Findings on evaluation consistency across generative AI tools

In Table 6, within the framework of the AI-as-Evaluator approach, the level of consistency among the evaluations conducted by the generative AI tools is analysed through standard deviation values, and the distribution of differences across the evaluator tools is presented.

Table 6. **Standard deviation values of strategies by evaluation criteria**

	Volatility strategy					Uncertainty strategy				
	I	F	A	R	M	I	F	A	R	M
ChatGPT	0,58	0,00	0,58	0,58	0,58	0,00	1,15	0,58	1,15	0,58
Gemini	0,58	0,58	0,58	1,00	1,15	1,15	1,15	1,00	0,00	0,58
DeepSeek	0,58	1,00	0,58	1,53	1,15	0,58	1,00	0,58	1,00	1,00
Claude	0,58	0,58	0,00	0,58	0,00	1,53	1,00	0,58	0,58	1,00
	Complexity strategy					**Ambiguity strategy**				
	I	**F**	**A**	**R**	**M**	**I**	**F**	**A**	**R**	**M**
ChatGPT	0,00	0,58	1,15	0,58	1,00	0,58	1,00	0,58	0,58	0,58
Gemini	1,73	2,52	1,00	1,00	1,00	0,00	0,00	0,58	0,58	0,58
DeepSeek	0,58	1,00	0,58	0,58	1,00	0,58	0,58	0,58	0,58	0,58
Claude	1,00	1,00	0,58	1,15	1,15	1,53	0,58	0,00	0,00	0,58

I: innovation and creativity, F: feasibility, A: agility and adaptability, R: risk level, M: market alignment

Source: own study

An examination of Table 6 shows that, except for some combinations, the standard deviation values by evaluation criterion are generally low. The fact that most standard deviation values cluster in the 0.58-1.15 range indicates that different evaluators tend to exhibit similar scoring patterns in many cases. The 0.00 values observed in certain strategy-criterion combinations, in turn, demonstrate that a high level of agreement can emerge among evaluators. However, the areas where consistency weakens most markedly are concentrated in the Complexity strategies. In particular, the elevated standard deviations for the Feasibility (F) criterion in these strategies (e.g., F: 2.52 for the Gemini strategy) suggest that, under conditions of complexity, the feasibility of strategy recommendations is interpreted in more divergent ways by the evaluators. Similarly, for the Innovation and Creativity (I: 1.73) criterion, and to some extent for the Risk (R: 1.53) criterion, evaluator divergence is observed to increase in certain strategy-criterion pairings.

At the criterion level, the fact that the Agility and Adaptability (A) scores have lower standard deviation values (0.00-1.15) for most strategies indicates that this criterion is scored with a more shared reference frame among evaluators. By contrast, it is observed that evaluator consistency systematically weakens for the Feasibility (F) criterion in Uncertainty strategies (1.00-1.15) and for the Risk (R) criterion in Volatility strategies (0.58-1.53). In Complexity strategies, the Innovation and Creativity (I) (1.73) and Feasibility (F) (2.52) criteria stand out as those with the highest deviations. These findings show that, although there is a general ground of consistency in the AI-as-Evaluator approach, as complexity and uncertainty increase-particularly for practice-oriented criteria-the divergence in scoring becomes more pronounced.

In summary, the findings of this study show that VUCA components significantly influence both the quality of generative AI-supported strategy outputs and the consistency of the evaluation processes applied to these outputs. In the Discussion section, these differences are examined in the context of the relevant theoretical frameworks, and the implications of the findings for the literature and for future research are assessed.

5. Discussion

The first finding of the study-that the raw evaluation scores reported in Table 3 are not uniform across strategies and criteria but display variation-is considered important in two respects. First, it shows that ignoring the components of VUCA and treating the concept as a single, undifferentiated uncertainty environment can be analytically misleading. This is consistent with the view that each component entails a different type of problem and therefore requires a distinct strategic response (Bennett & Lemoine, 2014a). Second, it is observed that generative AI tools, even when prompted with the same standardized prompt, do not produce generic outputs but instead develop strategies specific to the VUCA components, and that this diversity is maintained in the strategy evaluations as well. From the perspective of Information Processing Theory, this diversity indicates that the information processing requirements created by different VUCA scenarios do not align to the same extent, in every scenario, with the information processing capacity of the generative AI tools that both formulate and evaluate strategies (Biloslavo et al., 2025; Galbraith, 1974; Tushman & Nadler, 1978). Since VUCA components structure the nature of uncertainty in different ways, it is to be expected that, even when the same evaluation criteria are used, strategies will display varying levels of fit with these criteria. In this context, Table 3 provides strong evidence that the diversity observed in the mean and standard deviation values in the subsequent tables is not random but may reflect systematic divergences arising from the differing information processing demands of the VUCA components.

In the criterion-based mean results, the most striking point is that, as shown in Table 4, while the overall performance averages are relatively high for Volatility and Uncertainty strategies, performance declines for all tools in Complexity strategies, with a particularly downward clustering in the "Feasibility" (F) criterion. The VUCA literature describes volatility in terms of rapid fluctuations and shocks, and uncertainty in terms of weakened probability distributions and reduced predictive power, whereas it associates complexity with dense interdependencies among numerous factors and the loss of transparency in causal chains (Bennett & Lemoine, 2014a). This distinction in the literature appears critical for explaining the possible mechanism behind the observed finding. Volatility/Uncertainty conditions may be more conducive to generating more sequential and linear responses within a "fast decision" and "risk management" frame, whereas Complexity conditions, due to mutually influencing subsystems, make it more difficult to translate strategy into a concrete logic of action (feasibility).

Information Processing Theory posits that as uncertainty and interdependence increase, the amount of information required by decision makers rises, and that appropriate strategic responses must be supported by concrete mechanisms that either reduce this information need or enhance information processing capacity (Galbraith, 1974; Tushman & Nadler, 1978). In this study, the low "Feasibility" scores observed for Complexity strategies indicate that generative AI tools can produce persuasive answers at the strategic level to the question of "what should be done," but remain more limited in transforming these answers into context-specific and implementable action steps at the "how to do it" level. The bounded rationality perspective, on the other hand, suggests that under cognitive constraints, decision makers search not for the "best" options but for those that are "good enough," and that under complex conditions they tend to rely more on simpler options (Simon, 1955). The lower feasibility scores in Complexity strategies can be explained by the tendency of the tools to leave strategies mostly at the level of general principles and their difficulty in turning these recommendations into a more concrete implementation framework. Therefore, this finding points less to a "quality" difference between tools and more to the fact that the complexity component itself makes it harder to translate a strategy into an implementable structure. On the other hand, the relatively high and stable appearance of the "Agility and Adaptability"

(A) criterion across strategies suggests that generative AI tools are more inclined toward a "dynamic adaptation" discourse in strategy development. This finding is also consistent with the dynamic capabilities literature, which emphasizes that agility and the ability to reconfigure can be an important source of competitive advantage in rapidly changing environments (Teece et al., 1997).

The fact that the mean scores of the strategies generated by the generative AI tools are relatively close to each other in Volatility and Uncertainty strategies, but more divergent in Complexity strategies, is also theoretically meaningful. Eisenhardt (1989) showed that in high-velocity environments, fast decision-making teams use more real-time information, generate more alternatives, and manage conflict more effectively. Since Complexity requires integration and systems thinking rather than speed, it is plausible that differences in the reasoning styles of generative AI tools become more visible in this context. This finding indicates that the criteria for what constitutes a "good strategy" can vary across VUCA strategies and that differences between tools become more pronounced in certain types of strategies.

The fact that the standard deviation values emerging within the AI-as-Evaluator approach and reported in Table 5 are mostly clustered in the 0.58-1.15 range suggests that generative AI-based evaluators tend to display similar scoring patterns in many cases. This is broadly consistent with the LLM-as-a-judge literature, which argues that powerful generative AI tools can achieve high alignment with human preferences in certain tasks and can be used as scalable evaluation instruments (Zheng et al., 2023). At the same time, the same literature points to systematic risks in evaluator models, such as position bias, length bias, and various prompt sensitivities (Shi et al., 2024). Therefore, while the generally "high consistency" observed in this study's findings indicates that the AI-as-Evaluator approach is promising, it should not be overlooked that this consistency becomes fragile depending on the specific context and criteria under consideration. In particular, the higher standard deviation values observed for the Feasibility (2.52) and Innovation and Creativity (1.73) criteria in Complexity strategies suggest that the evaluator tools rely on different assumptions when interpreting these two criteria. Likewise, as shown in Table 5, the fact that the scores for the "Agility and Adaptability" criterion exhibit generally lower standard deviation values indicates that this criterion is evaluated by the evaluators using a more shared language and template. This is thought to stem from the tendency of generative AI tools to frequently resort to similar discursive patterns when producing outputs. By contrast, criteria such as "Feasibility," which focus on the extent to which strategies can be implemented under concrete organizational conditions, may be more sensitive to missing information in the scenarios and more exposed to the evaluators' implicit assumptions, thereby leading to greater divergence.

When the findings are considered as a whole, they show that VUCA components affect not only the nature of strategy formulation but also the consistency of strategy evaluation. This result makes the discriminating power of VUCA across its components methodologically visible (Bennett & Lemoine, 2014a). The study also demonstrates that the performance of generative AI–supported strategy development is relatively strong on common and abstract strategic themes such as "agility," whereas under conditions of Complexity it can display a more fragile performance on criteria such as feasibility, which are more sensitive to resources and organizational arrangements. This finding supports a more in-depth discussion of the notion that generative AI tools do not replace managerial decision makers but rather assume an enabling role that supports their strategic thinking and evaluation capacities (Raisch & Krakowski, 2021).

6. Conclusion

This study demonstrates that strategy performance developed in the context of different VUCA scenarios varies systematically (in particular, that the complexity component causes a marked decrease in the "implementability" scores), and that this variation is related not only to the content of the strategies but also to the evaluation logics of the generative AI tools. This result indicates that it is misleading to treat strategic correctness as a single, context-independent criterion. Therefore, it is evidenced that, particularly in highly uncertain environments, strategy performance gains meaning through multiple, context-based evaluations.

From a theoretical perspective, the study extends information processing theory and the bounded rationality approach in the context of strategy development supported by generative AI. It also shows that generative AI tools should be conceptualized not as actors that replace strategic decision makers but, with their relatively high and stable "agility and adaptability" performance across all scenarios, as complementary structures that restructure strategic reasoning processes and expand cognitive boundaries in VUCA contexts. Methodologically, the study demonstrated that the Artificial Intelligence-Based Evaluation (AI-as-Evaluator) approach offers a practical framework for analysing context-sensitive strategic reasoning patterns - without claiming absolute accuracy - by leveraging the general scoring consistency and consistent evaluation trends observed among evaluator models.

This study reveals that each component of VUCA creates distinct challenges for strategic reasoning and that these challenges lead to systematic differences in the strategy recommendations produced by generative AI tools. This extends the core assumption of information processing theory regarding the fit between environmental uncertainty and information processing capacity to explicitly include AI. At the same time, within the framework of the bounded rationality approach, it shows that generative AI tools provide an information processing capacity that can alleviate the cognitive limitations of human decision makers.

In addition, by enabling strategy recommendations to be scored through blind and multiple generative AI evaluators within the AI-as-Evaluator approach, this study demonstrates that the approach offers an applicable framework for strategy evaluation processes. However, the fact that the consistency among generative AI evaluators varies across VUCA components indicates that strategic correctness is not a single, objective criterion, but rather a context-dependent, perceptual one. This finding highlights the limits of the established "one best strategy" view in the strategy literature and clearly shows why this view falls short in explaining VUCA contexts.

The findings indicate that generative AI tools are more appropriately regarded not as substitutes for strategic decision makers, but as "strategic thinking and decision-support partners" that generate alternative strategic frames under different VUCA conditions and reduce the cognitive load on decision makers. However, while managers position the generative AI as an agility-focused partner, in situations involving high complexity, they must ensure that the strategic responses offered by these tools are always reviewed by humans in terms of operational feasibility. Moreover, in environments where ambiguity and complexity are predominant, instead of relying on a single generative AI output, comparatively evaluating the strategies developed by different models stands out as a management practice that can improve decision quality. In this respect, the study provides a context-sensitive and methodologically grounded contribution to both theoretical debates and managerial practices concerning generative AI-supported strategy development. On the other hand, when putting generative AI-supported strategy development and evaluation processes into practice, it should not be overlooked that access to these technologies is not distributed equally across organizations and that, at the implementation stage, the choice of tools is often determined by the institution's limited resources and access opportunities.

The findings of this study should be interpreted within the scope of certain methodological and contextual limitations. First, the strategy development process was conducted in a single-shot manner for each scenario-generative AI tool pair. Due to the possibility that generative AI outputs may vary over time and with changing system conditions, this design limits the generalizability of the strategy recommendations obtained. Second, although the adoption of the AI-as-Evaluator approach in the evaluation of strategy performance provides a robust framework in terms of methodological consistency and blindness, the exclusion of human expert assessments from the process is regarded as a factor that constrains the validity of the findings.

Future research may adopt multiple designs that involve more than a single run in the strategy development stage to overcome these limitations. In addition, the generalizability of the findings can be tested by employing different evaluation criteria in the strategy assessment stage and by using hybrid evaluation models in which not only generative AI tools, but also human experts and generative AI evaluators are involved together. In addition to the scenarios in which the VUCA components are addressed individually, examining complex scenarios where these dimensions emerge simultaneously and interactively, as they do in the real world, offers an important area of research that can reveal the limits of generative AI-assisted strategic reasoning in greater depth. Finally, future research should focus on empirically testing the performance of AI models with more advanced reasoning abilities - aimed at addressing AI's weakness in applying strategies in "complex" situations - and the impact of human-machine collaboration on strategic success.

Authors' contribution

E.Y.: article conception, research methods applied, conducting the research, data collection, analysis and interpretation of results. **C.Y.:** theoretical content of the article, draft manuscript preparation, conducting the research.

Declaration of Generative AI and AI-assisted technologies in the writing process

In this study, GenAI tools were used as a supportive tool at specific stages of the research process. The ScholarGPT tool was used to generate scenarios representing the VUCA components. Based on these scenarios, strategy recommendations have been developed using ChatGPT, Gemini, DeepSeek, and Claude. In addition, the same GenAI tools were used to evaluate the strategies developed. Although GenAI tools were used as part of the research design, the results obtained were critically examined and validated by the authors, and the final interpretations were formulated by the researchers. The authors of the GenAI tools have not been included as authors, and all scientific and ethical responsibility for the study rests with the authors.

References

Almutairi, W., & Almatrodi, I. (2025). Trust under bounded rationality: Exploring human-ai interaction in decision-making through large language models. *SAGE Open, 15*(4). https://doi.org/10.1177/21582440251380135

Bader, B., Schuster, T., Bader, A. K., & Shaffer, M. (2019). The dark side of expatriation: Dysfunctional relationships, expatriate crises, prejudice, and a VUCA world. *Journal of Global Mobility, 7*(2), 126-136. https://doi.org/10.1108/jgm-06-2019-070

Baran, B. E., & Woznyj, H. M. (2021). Managing VUCA: The human dynamics of agility. *Organizational Dynamics, 50*(2), 100787.

Benhür Aktürk, E. (2025). Strategic decision-making and artificial intelligence: Exploring the impact of AI applications on decision precision and risk mitigation. *Journal of Society Research, 22*(4), 580-592. https://doi.org/10.26466/opusjsr.1666747

Bennett, N., & Lemoine, G. J. (2014a). What a difference a word makes: Understanding threats to performance in a VUCA world. *Business Horizons, 57*(3), 311-317. https://doi.org/10.1016/j.bushor.2014.01.001

Bennett, N., & Lemoine, G. J. (2014b). What VUCA really means for you. *Harvard Business Review, 92*(1/2). Retrieved February 9, 2026 from https://hbr.org/2014/01/what-vuca-really-means-for-you

Bessa, G., & Barbosa, B. (2025). Integrating artificial intelligence into scenario analysis: a validated framework for strategic planning under economic uncertainty. *Global Economics Research, 1*(2), 100007. https://doi.org/10.1016/j.ecores.2025.100007

Biloslavo, R., Edgar, D., Aydin, E., & Bulut, C. (2025). Artificial intelligence (AI) and strategic planning process within VUCA environments: A research agenda and guidelines. *Management Decision, 63*(10), 3599-3624. https://doi.org/10.1108/MD-10-2023-1944

Cernega, A., Nicolescu, D. N., Imre, M. M., Totan, A. R., Arsene, A. L., Şerban, R. S., Perpelea, A. C., Nedea, M. I., & Piţuru, S. M. (2024). Volatility, uncertainty, complexity, and ambiguity (VUCA) in healthcare. *Healthcare,* 12, 773. https://doi.org/10.3390/healthcare12070773

Csaszar, F. A., Ketkar, H., & Kim, H. (2024). Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors. *Strategy Science, 9*(4), 322-345. https://doi.org/10.1287/stsc.2024.019

de Mesentier Silva, F., Lee, S., Togelius, J., & Nealen, A. (2017). AI as evaluator: Search driven playtesting of modern board games. In N. R. Sturtevant, A. Isaksen, J. Togelius, & J. Zhu (Eds.), What's next for AI in games? Papers from the 2017 AAAI Workshop, San Francisco, California, USA, February 4, 2017 (Vol. WS-17-15, pp. 959–959). Association for the Advancement of Artificial Intelligence Press.

Doz, Y. L., & Kosonen, M. (2010). Embedding strategic agility: A leadership agenda for accelerating business model renewal. *Long Range Planning, 43*(2-3), 370-382. https://doi.org/10.1016/j.lrp.2009.07.006

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data - Evolution, challenges and research agenda. *International Journal of Information Management,* 48, 63-71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal, 32*(3), 543-576. https://doi.org/10.2307/256434

Fridgeirsson, T. V., Ingason, H. T., Jonasson, H. I., & Kristjansdottir, B. H. (2021). The vucality of projects: A new approach to assess a project risk in a complex world. *Sustainability, 13*(7), 3808. https://doi.org/10.3390/su13073808

Galbraith, J. R. (1974). Organization design: An information processing view. *Interfaces, 4*(3), 28-36. https://doi.org/10.1287/inte.4.3.28

Galbraith, J. R. (1977). *Organization design.* Addison-Wesley Publishing Company.

Gu, J., Jiang, X., Shi, Z., Tan, H., Zhai, X., Xu, C., Li, W., Shen, Y., Ma, S., Liu, H., Wang, S., Zhang, K., Lin, Z., Zhang, B., Ni, L., Gao, W., Wang, Y., & Guo, J. (2025). A survey on LLM-as-a-judge. *arXiv.* https://doi.org/10.48550/arXiv.2411.15594

Gürsoy, A. Ö., & Şavk, S. (2024). From scriptor to promptor: An evaluation of the status of authorship, authenticity, and creativity in light of the use of artificial intelligence in screenwriting. *ARTS,* 12, 57-82. https://doi.org/ 10.46372/arts.1482636

Hastings, S. (1996). A strategy evaluation model for management. *Management Decision, 34*(1), 25-34. https://doi.org/ 10.1108/00251749610106945

Horney, N., Pasmore, B., & O'Shea, T. (2010). Leadership agility: A business imperative for a VUCA world. *People & Strategy, 33*(4), 32-38.

Jacob, S. (2025). Artificial intelligence and the future of evaluation: From augmented to automated evaluation. *Digital Government: Research and Practice, 6*(1), 1-10. https://doi.org/10.1145/3696009

Kesgin, K. (2025). Yapay zekâ tabanlı karar destek sistemlerinde teknik, etik ve yönetişimsel yaklaşımlar: Çok sektörlü bir derleme çalışması [Technical, Ethical, And Governance Approaches in Artificial Intelligence-Based Decision Support Systems: A Multi-Sectoral Review]. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 9*(1), 29-48. https://doi.org/10.33461/uybisbbd.1691218

Kinsinger, P. & Walch, K. (2012). *Living and leading in a VUCA world.* Thunderbird University.

Kitsios, F., & Kamariotou, M. (2021). Artificial intelligence and business strategy towards digital transformation: A research agenda. *Sustainability, 13*(4), 2025. https://doi.org/10.3390/su13042025

Laskar, M. T. R., Jahan, I., Dolatabadi, E., Peng, C., Hoque, E., & Huang, J. X. (2025). Improving automatic evaluation of large language models (LLMs) in biomedical relation extraction via LLMs-as-the-judge. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers),* 25483-25497. Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.acl-long.1238

Lawrence, K. (2013). *Developing leaders in a VUCA environment.* UNC Executive Development. https://www.emergingrnleader.com/wp-content/uploads/2013/02/developing-leaders-in-a-vuca-environ ment.pdf

Li, H., Dong, Q., Chen, J., Su, H., Zhou, Y., Ai, Q., Ye, Z., & Liu, Y. (2024). LLMs-as-Judges: A comprehensive survey on LLM-based evaluation methods. *arXiv.* https://doi.org/10.48550/arXiv.2412.05579

Liu, Y. C., Huang, C. M. K., Chang, Y. S., Lin, H. M., & Chen, P. L. (2023). An integrative model of information processing and contextual factors on exploring information systems outsourcing success. *Asia Pacific Management Review, 28*(3), 327-335. https://doi.org/10.1016/j.apmrv.2022.12.001

López-Solís, O., Luzuriaga-Jaramillo, A., Bedoya-Jara, M., Naranjo-Santamaría, J., Bonilla-Jurado, D., & Acosta-Vargas, P. (2025). Effect of generative artificial intelligence on strategic decision-making in entrepreneurial business initiatives: A systematic literature review. *Administrative Sciences, 15*(2), 66. https://doi.org/ 10.3390/admsci15020066

Luo, B. N., & Donaldson, L. (2013). Misfits in organization design information processing as a compensatory mechanism. *Journal of Organization Design, 2*(1), 2-10. https://doi.org/10.7146/jod.7359

Mack, O., & Khare, A. (2016). Perspectives on a VUCA world. In: Mack, O., Khare, A., Krämer, A., Burgartz, T. (Eds.), *Managing in a VUCA World.* Springer, Cham. https://doi.org/10.1007/978-3-319-16889-0_1

Oral, D. Ö. (2024). *Yapay zekanın senaryo üretiminde işlevselliği: ChatGPT-4o* örneği [Functionality of artificial intelligence in scriptwriting: The example of ChatGPT-4o]. [Master's thesis, Karabük University]. YÖK National Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=UjlM15wKZGQW6TLC0pv Ct1sLNo-gXUVlbnQj-bRaiQIwRBZHx8Ek7p158YKEloL

Perifanis, N.-A., & Kitsios, F. (2023). Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. *Information, 14*(2), 85. https://doi.org/10.3390/info14020085

Pittenger, L. M., Glassman, A. M., Mumbower, S., Merritt, D. M., & Bollenback, D. (2023). Bounded rationality: Managerial decision-making and data. *Journal of Computer Information Systems, 63*(4), 890-903. https://doi.org/ 10.1080/08874417.2022.2111380

Raisch, S., & Krakowski, S. (2021). Artifcial intelligence and management: The automation-augmentation paradox. *Academy of Management Review, 46*(1), 192-210. https://doi.org/10.5465/amr.2018.0072

Rumelt, R. P. (1979). Evaluation of strategy: Theory and models. In Schendel, D. E. & Hofer, C. W. (Eds.), *Strategic management: A new view of business policy and planning* (196-212). Little, Brown.

Rzepczynski, M. (2025). *Clarifying the assessment of risk: VUCA (Volatility, Uncertainty, Complexity, and Ambiguity).* AMPHI Research and Trading. https://doi.org/10.2139/ssrn.5217110

Shet, S. V. (2024). A VUCA-ready workforce: Exploring employee competencies and learning and development implications. *Personnel Review, 53*(3), 674-703. https://doi.org/10.1108/PR-10-2023-0873

Shi, L., Ma, C., Liang, W., Ma, W., & Vosoughi, S. (2024). Judging the judges: A systematic investigation of position bias in pairwise comparative assessments by LLMs. In P*roceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics* (pp. 292–314), Mumbai, India. The Asian Federation of Natural Language Processing and the Association for Computational Linguistics.

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review, 61*(4), 66-83. https://doi.org/10.1177/0008125619862257

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118. https://doi.org/10.2307/1884852

Simon, H. A. (1991). Bounded rationality and organizational learning. *Organization Science, 2*(1), 125–134.

Simon, H. A. (1997). *Administrative behavior* (4. Ed.). Free Press.

Steiner, G. A. (1969). *Top management planning.* Macmillan.

Sullivan, J. (2012). VUCA: The new normal for talent management and workforce planning. Retrieved Feb, 9, 2026 from https://www.ere.net/articles/vuca-the-new-normal-for-talent-management-and-workforce-planning

Syamsir, S., Saputra, N., & Mulia, R. A. (2025). Leadership agility in a VUCA world: A systematic review, conceptual insights, and research directions. *Cogent Business & Management, 12*(1). https://doi.org/10.1080/ 23311975.2025.2482022

Takemoto, K. (2024). The moral machine experiment on large language models. *Royal Society Open Science, 11*(2), 231393. https://doi.org/10.1098/rsos.231393

Taskan, B., Junça-Silva, A., & Caetano, A. (2022). Clarifying the conceptual map of VUCA: A systematic review. *International Journal of Organizational Analysis, 30*(7), 196-217. https://doi.org/10.1108/IJOA-02-2022-3136

Tavana, M., & Banerjee, S. (1995). Strategic assessment model (SAM): A multiple criteria decision support system for evaluation of strategic alternatives. *Decision Sciences, 26*(1), 119-143. https://doi.org/10.1111/j.1540-5915.1995.tb00840.x

Teece, D. J., Peteraf, M. A., & Leih, S. (2016). Dynamic capabilities and organizational agility: Risk, uncertainty and entrepreneurial management in the innovation economy. *California Management Review, 58*(4), 13-35. https://doi.org/10.1525/cmr.2016.58.4.13

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

Thorén, K., & Vendel, M. (2019). Backcasting as a strategic management tool for meeting VUCA challenges. *Journal of Strategy and Management, 12*(2), 298-312. https://doi.org/10.1108/JSMA-10-2017-0072

Tushman, M. L., & Nadler, D. A. (1978). Information processing as an integrating concept in organizational design. *Academy of Management Review, 3*(3), 613-624. https://doi.org/10.2307/257550

Vargas-Hernández, J. G., & Pérez Ortega, R. (2019). Bounded rationality in decision-making. *MOJ Current Research and Review, 2*(1), 1-8. https://doi.org/10.15406/mojcrr.2019.02.00047

Wei, Y. M. (2025). A hybrid multi-criteria decision-making framework for the strategic evaluation of business development models. *Information, 16*(6), 454. https://doi.org/10.3390/info16060454

Winkler, P., Kretschmer, J., & Wamprechtsamer, P. (2025). Navigating through digitalization challenges in strategic communication: Introducing the VUCA radar. *Journal of Communication Management, 29*(2), 240-257. https://doi.org/10.1108/JCOM-11-2023-0119

Wulf, T., Meissner, P., Brands, C., & Stubner, S. (2012). Scenario-based strategic planning: A new approach to coping with uncertainty. In: Schwenker, B., Wulf, T. (Eds.), *Scenario-based Strategic Planning.* Roland Berger School of Strategy and Economics. Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-02875-6_3

Yang, Y., Rahman, A. A., Abdan, K., Aziz Abdul, Y., & Li, Y. (2025). The application of organizational information processing theory in supply chain management strategy research: A bibliometric review. *Corporate & Business Strategy Review, 6*(1), 378-391. https://doi.org/10.22495/cbsrv6i1siart14

Yarger, H. R. (2006). *Strategic theory for the 21st century: The little book on big strategy.* US Army War College Press.

Yawson, R. M., & Goryunova, E. (2025). Nested complexity: A conceptual framework for leveraging ai for sustainable organizations and human resource development. *Advances in Developing Human Resources, 27*(2-3), 91-123. https://doi.org/10.1177/15234223251335908

Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E., Zhang, H., Gonzalez, J., & Stoica, I. (2023). Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, & S. Levine (Eds.), *Advances in Neural Information Processing Systems* (Vol. 36, pp. 46595–46623). Curran Associates, Inc.