Occam's Razor is a Theory: Why Cause-and-Effect Performance Links Form Parsimonious Mental Models of Complex Strategic Environments

Madison Singell

"He who loves practice without theory is like the sailor who boards a ship without a rudder."
~Leonardo da Vinci

Most Updated Version of Paper: https://mhsingell.github.io/files/Singell_JMP_Current_2025.pdf

1. INTRODUCTION

The theory-based view of strategy proposes that theories – mental models that specifically represent cause-and-effect relationships between strategic choices — find successful strategies (Felin et al., 2024; Camuffo et al., 2024). This theory-based view argues that specifying a causal pathway linking strategic choices to performance helps decision-makers more effectively experiment and learn, allowing them to select more successful strategies over time (Camuffo et al., 2024; Valentine et al., 2024; Sorenson, 2024; Ehrig & Schmidt, 2022). Yet, while there is some evidence that theories can improve strategic search (Camuffo et al., 2020), there is no model that explains why these mental models with cause-and-effect performance links would find better strategies than the mental models used in the large body of work that preceded this theory-based turn.

Following the seminal work of Porac, Thomas, and Baden-Fuller (1989), a stream of cognitive strategy research has long argued that mental models help decision-makers find more successful strategies (see Kaplan, 2011 for a review). One of the core mechanisms through which these mental models improve strategic search is through simplifying the complex strategic environment (Gavetti & Levinthal, 2000; Bingham & Eisenhardt, 2011). By focusing on a subset of strategic choices, decision-makers can more easily search the environment, better attribute performance variation, and avoid getting stuck on poor strategies (Csaszar & Levinthal, 2016).

The major difference between the mental models in this cognitive strategy research and the mental models in the theory-based view is the type of performance relationship they represent between strategic choices. In the cognitive strategy research tradition, mental models have implicitly associative performance links. These mental models, which I will refer to as associative mental models, represent the performance of two strategic choices as related (even causally related), but do not specify the direction of the causality (see Gavetti & Levinthal, 2000; Csaszar & Levinthal, 2016; Csaszar & Ostler, 2020).[1] In the theory-based view of strategy, mental models have cause-and-effect performance links. These mental models, which I will refer to as causal mental models or theories, represent the selection of one strategic choice as the performance antecedent of another (see Camuffo et al., 2024; Sorenson, 2024).

This difference in performance link type of mental models is likely to impact the performance of strategies selected by them because it affects how accurately and simply the mental model represents the strategic environment. Previous research suggests that both simple and accurate mental models find high-performing strategies, because simple models are easier to understand and accurate models are right (Gavetti & Levinthal, 2000; Gary & Wood, 2010). When the strategic environment is simple, mental models can be both simple and accurate. However, as complexity in the strategic environment increases, decision-makers using mental models to find strategies face a trade-off between simply and accurately representing it. In these complex environments, cognitive strategy research shows that associative mental models improve strategy selection through representing fewer strategic choices (Csaszar & Levinthal, 2016.

---

[1] Contingency functions ($f(x)$) which specify the interdependence of two variables (x_1 and x_2) as an interaction term ($x_1 x_2$) such that $f(x) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$, can always be differentiated in terms of both variables in the interaction ($\frac{\partial f}{\partial x_1} = \beta_1 + \beta_3 x_2$ and $\frac{\partial f}{\partial x_2} = \beta_2 + \beta_3 x_1$). Thus, these functions, by allowing both variables in the interaction to vary as a function of the other, do not specify causal direction of dependence between them (for example Csaszar & Ostler, 2020).

While this prior work suggests that simply representing complex strategic environments is likely to generate more performant strategies, it tends to explore one facet of simplification, that of considering fewer strategic choices (see Martignoni et al., 2016 for an exception). However, mental models are made up of two components: a set of strategic choices and a set of performance relationships between those choices. Thus, another way a mental model could simplify the strategic environment is through representing fewer performance links between strategic choices, which is what the cause-and-effect performance links of theories do.

I propose that the reason theories are likely to outperform associative mental models in selecting strategies is because they further simplify the strategic environment by representing fewer performance links between choices. Because theories specify a causal direction, they order the decision-maker's strategic choices, significantly reducing the number of conditional performance dependencies the decision-maker considers. This simplified consideration of performance links between strategic choices allows the decision-maker to better use performance feedback to inform future decisions (see credit assignment, Denrell et al., 2004), which leads to the selection of more performant strategies over time.

Using a model that compares how decision-makers employ different mental models to find strategies in increasingly complex environments, I find that theories handle the trade-off between accurately and simply representing performance links better than associative mental models, finding more successful strategies. Specifically, as the number of strategic choices in the environment increases, theories remain both more accurate and simpler representations of the environment when compared to associative mental models. And as the number of performance links between strategic choices in the environment grows, generating a trade-off between

accuracy and simplicity, theories make the more performant choice of representing these strategic environments more simply.

My work contributes to the growing stream of research on the theory-based view of strategy by suggesting that one of the benefits of holding a theory is that its cause-and-effect performance links more parsimoniously represent complex strategic environments, consistently finding more successful strategies than the equivalent associative mental model. Thus, the cause-and-effect performance links of a theory make it more effective at forming a strategy than an associative mental model of performance, even if that theory is less accurate. Overall, I suggest that theories are a powerful tool for organizational decision-makers looking to form successful strategies.

2. CHOOSING A STRATEGY WITH DIFFERENT MENTAL MODELS: AN EXAMPLE

To illustrate the difference between strategic decision-making using a theory versus an associative mental model, I use the example of a founder of a health tech startup trying to decide on three key features of her AI model built to detect early signs of cancer: data, learning, and evaluation. The key difference in decision-making is that theories prompt sequential decisions, while associative mental models prompt a set of simultaneous strategic choices.

In order to build a high-valuation startup around an AI model that detects cancer, the startup founder has many interdependent choices to consider. For example, the founder needs to decide which data her AI model should be trained on, where she could either choose a dataset with a lot of breadth (a few indicators for many different types of cancer) or depth (many indicators for a single cancer type, i.e. breast cancer). A model that uses a wide breadth of data will make the startup more marketable to a wider set of potential customers, while a model that

has depth of data is likely to make more accurate predictions for a particular type of cancer, making it more valuable to a smaller set of customers.

However, the impact of the choice of data on the valuation of the start-up is also dependent on the other model features that the founder selects, such as the way the model learns and the way that model success is evaluated. In selecting the way the AI model learns, the founder considers whether to train her model in a supervised or unsupervised way, where supervised learning requires that the data have labeled outcomes and unsupervised learning does not. The cost of labeling outcomes on a broader dataset may be higher and the benefit of unsupervised learning on a smaller dataset may be lower, thus the type of learning of the model is likely to interact with the type of data chosen to produce different model results and different valuations of the startup.

Finally, the founder considers the evaluation metric for the cancer predictions of her model, where it is likely that different datasets and learning methods will generate differential success by evaluation metric. Two of the major evaluation metrics in AI model predictions are precision, which measures the accuracy of positive predictions, and recall, which measures the model's ability to identify all positive cases. Since precision minimizes the number of people the model incorrectly identifies as having cancer and recall maximizes the number of people the model correctly identifies as having cancer, both outcomes may be desirable for the startup and its valuation. However, the ability of the model to have high precision and recall is likely a function of the choice of data and learning that the startup founder selects.

Thus, for a health tech startup founder building an AI model to detect cancer, it will be important for her to consider how the data (breadth vs. depth), learning method (supervised vs. unsupervised), and evaluation metric (precision vs. recall) selected interact to generate different

valuations for the startup. Figure 1 shows the strategic choice interdependencies for this startup founder building an AI model for cancer detection and how their combination impacts the overall valuation of the health tech startup.

FIGURE 1
Strategic Environment of the Performance Links
Between the AI Model for Cancer Detection and the Health Tech Startup's Valuation
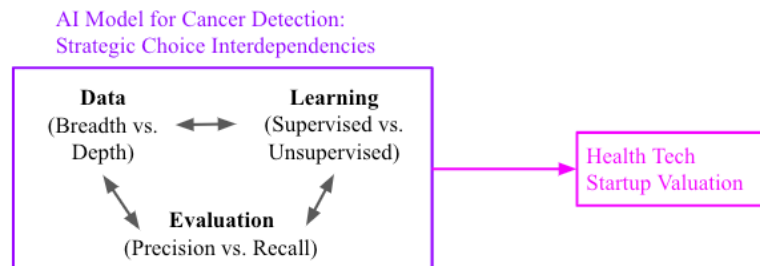
To handle the complex links between the AI model's features, the health tech startup's founder could use a mental model to guide her strategy selection. Consistent with previous research, her mental model could consist of associative performance links between a subset of the model features, improving her decision-making by focusing her attention on a subset of strategic choices (Gavetti & Levinthal, 2000; Csaszar & Levinthal, 2016). Figure 2 shows this associative type of mental model, where the founder focuses on choosing the data and learning method for her model simultaneously, while deciding on the evaluation metric independently.

FIGURE 2
Mental Model with Associative Performance Links Selecting the AI Model Strategy
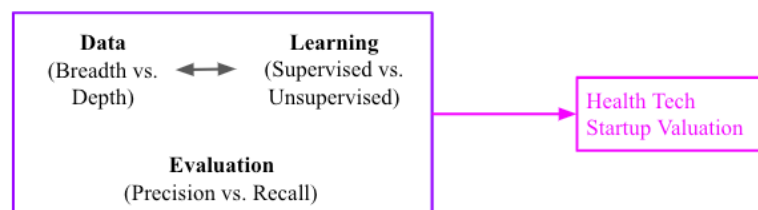
Alternatively, the decision-maker could use a theory — a mental model with cause-and-effect performance links— to choose the AI model's features. As shown in Figure 3, this model

also focuses on the links between a subset of the strategic choices: the data and the learning type. However, unlike the associative model, the links in the theory have a causal direction, where the selection of the data type is seen as a performance antecedent of the learning type. This ordering of strategic choices in the theory prompts the decision-maker to first choose the data type (breadth vs. depth), and then choose the learning type of the model conditional on the data type selection. Like in the associative model, the evaluation metric is chosen independently.

FIGURE 3
Mental Model with Cause-and-Effect Performance Links Selecting the AI Model Strategy
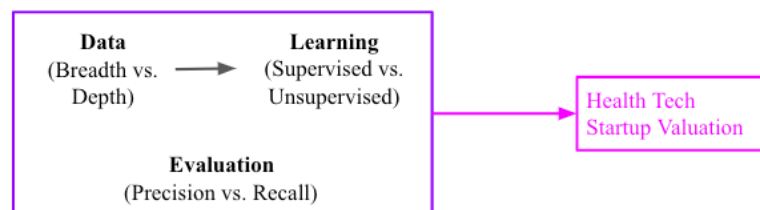
The above example highlights how the way a decision-maker chooses a strategy depends on the performance link type (associative vs. cause-and-effect) in her mental model. Cause-and-effect performance links prompt sequential decisions, while associative performance links prompt simultaneous choice. Because these two types of mental models suggest different strategic decision-making processes, they are likely to interpret performance feedback differently, resulting in different strategies and performance outcomes (Denrell et al., 2004; Martignoni et al., 2016). Next, I use a model to explore how mental models with associative and cause-and-effect performance links affect strategy selection and performance.

3. MODEL

Adapting the canonical NK-landscape model (Levinthal, 1997; Gavetti & Levinthal, 2000), I find that mental models with cause-and-effect performance links (or theories) outperform associative models in finding successful strategies because they are more

parsimonious representations of the strategic environment. Table 1 below outlines the key parameters for the model.

TABLE 1: MODEL PARAMETERS

Object	Symbol	Description	Category of Variance
Time	T	The number of time periods, T = 100	System
Number of Simulations	S	The number of times to iterate through different landscapes, S = 1000	System
Number of Strategic Choices	N	Number of strategic choices in environment, N=10 unless noted	Strategic Environment
Interdependencies	K	Number of factors that the performance of each N choice is dependent on, K = 3 unless noted	Strategic Environment
Size of Mental Model (M)	M	The number of strategic choices in the mental model	Mental Model
Order	OR	A randomly selected order of length M of the form A→B→C→….. →M	Mental Model (for theory only)

3.1 Characterizing the Strategic Environment of Organizations as an NK-Landscape

Consistent with previous strategy research, I use the NK-landscape model to characterize the complex strategic environment of organizations (Levinthal, 1997; see Baumann et al., 2019 for a review). In this model, strategy selection is conceptualized as a decision-maker traversing a 'rugged' landscape, where the difficulty of finding successful strategies depends on the number of strategic choices (N) and the number of performance links between each strategic choice (K).

The NK-landscape model presents a particular type of problem that faces decision-makers selecting strategies in complex environments.

Often called an NP-hard or NP-complete problem in mathematics, the NK-landscape formalizes strategic environments that are so complex that past performance feedback cannot guide decision-makers to the best strategy (see strategy as an NP-hard problem: Siggelkow, 2011; Ganco & Hoetker, 2009; Hochba, 1997). As these environments grow larger (high N) and more "rugged" (high K), it becomes harder to understand how each choice affects performance, since similar strategies can lead to very different outcomes (Rivkin, 2000). Thus, decision-makers facing these complex strategic environments must find effective ways to use performance feedback to improve their strategic choices.

3.2 Mechanism: The Selection of Strategies by Mental Model Link Type

The performance link type of a decision-maker's mental model determines how the decision-maker uses previous performance feedback to select future strategies (Denrell et al., 2004). Expanding on the intuition of the motivating example, I build a model of the simultaneous and sequential choice prompted by mental models with associative and cause-and-effect performance links. This model shows that associative and causal mental models prompt decision-maker to consider different sets of conditional performance dependencies when analyzing prior performance data, which leads to the selection of different strategies, even when considering the same M strategic choices.

In associative mental models the direction of the performance dependence between strategic choices is not specified, instead the model suggests a bundle of M strategic choices that are related. In our motivating example, the startup founder considers the data and learning type as a bundle of choices (M=2). Work in managerial cognition and routines shows that when

choices are grouped together, the performance implications of the choices are often considered jointly (Nelson & Winter, 1982; Levinthal, 2000; Tripsas & Gavetti, 2000; Gavetti & Rivkin, 2007). This means that the founder, upon trying a dataset with depth and an unsupervised learning approach and receiving poor performance feedback, will conclude that this combination of model features does not perform well, and will select a different combination in future strategies.

To operationalize this bundling of choices in the model, the decision-maker considers the previous performance of all M strategic choices and selects the highest average performing combination. For the N-M strategic choices not included in the mental model she selects the highest performing value for each independently. If this process selects a strategy that has already been selected, the decision-maker randomly selects one of the N strategic choices to change. This approach is consistent with the operationalization of mental models in prior work (see Choi & Levinthal, 2023; Gavetti & Levinthal, 2000; Brusoni et al., 2007).

In a theory, the performance dependence between strategic choices is specified, where the theory orders the performance dependence such that strategic choices are antecedents and consequents of one another. In the motivating example, the startup founder's theory suggests that the data type is a performance antecedent of the learning type (M is still equal to 2). This ordering leads the decision-maker to use previous performance feedback to select the best possible data type (for example a broad dataset), and then to select the best possible learning type conditional on the data type being broad (see Camuffo et al., 2024). If she tries a broad dataset and a supervised learning approach and sees poor performance, she concludes that conditional on selecting a dataset with breadth, selecting supervised learning is a poor strategy (see Denrell et al., 2004 or Martignoni et al., 2016 for an example of sequential credit assignment).

To operationalize this sequential choice in the model, the decision-maker forms a theory where the performance links are randomly selected into an ordered linear chain of length M, i.e. A→B→…→M.[2] (footnote 2: This is a simple and common structure of cause-and-effect performance dependence in organizations, but there are other structures that are possible (see Pearl, 2009; Georgakopoulos et al., 1995; Gary & Wood, 2010)). Using this chain of causal performance dependence, the decision-maker first selects the best performing value for strategic choice A, then, conditional on the selection of strategic choice A, selects the most performant value for strategic choice B. She continues this pattern of strategic choice using conditional performance dependence until she runs out of performance data, after which she explores with the next choice, and leaves the remaining M choices at their previous maximal value. Like associative models, cause-and-effect models also have the decision-maker decide on the N-M strategic choices not included in the model independently based on prior performance feedback, and modify a random N item if the strategy has been tried before. Appendix A contains the pseudo-code and link to the full code for both the operationalization of strategic search using both an associative mental model and a theory.

While both mental model types engage in greedy search for a strategy based on prior performance feedback, the difference in the process of incorporating performance feedback between the link types leads the decision-maker to select different strategies. Figure 4 below shows the average proportion of shared strategies selected by a decision-maker using the same M strategic choices in an associative mental model vs. theory across 100 times periods and 1000 simulation runs (for strategic environment, N = 10, K = 3).
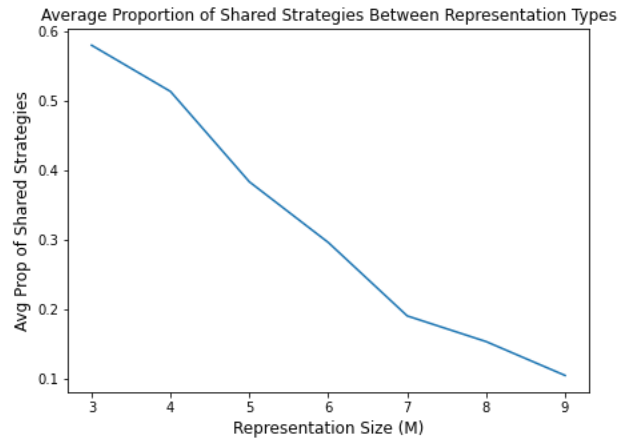
FIGURE 4

Figure 4 shows that the difference in link type of the mental models leads the decision-maker to select different strategies between 40 to 90 out of 100 times, depending on the size of the representation used. This significant difference is strategy selection is true even when these mental models searched the same strategic environment, with the same M strategic choices in the mental model, and started with the same initial strategy. Because mental models with associative vs. cause-and-effect performance links prompt decision-makers to incorporate performance feedback in either a simultaneous or sequential way, these mental models generate significant difference in the strategies that the decision-maker selects. To develop an understanding for why the strategies selected by theories might perform better than those selected by associative mental models, I next turn to how simply and accurately each type of mental model represents the underlying strategic environment.

3.3 The Impact of Mental Model Performance Link Type on Parsimonious Representation of the Strategic Environment

Mental models are representations of the strategic environment that help decision-makers select strategies. Previous research suggests that these representations can help decision-makers address the core challenge of correctly attributing strategic performance in complex

environments through simplifying the number of strategic choices that the decision-maker considers (Gavetti & Levinthal, 2000). However, mental models that accurately represent the strategic environment are also more likely to attribute strategic performance correctly (Gary & Wood, 2010). Below I model how the performance link type of mental models (cause-and-effect vs. associative) impacts the degree to which these mental models accurately and simply represent the strategic environment, which varies by the complexity of this environment.

In simple environments, mental models can be both accurate and simple, and in these cases, I find that theories are more accurate and simpler representations than associative models. As the number of strategic choices increases (N), theories remain both simpler and more accurate representations. But when complexity comes from a greater number of performance links between choices (K), theories are simpler, while associative mental models are more accurate.

3.3.1 The Effect of a Mental Model's Performance Link Type on the Simplicity of the Representation

To model how simply a mental model represents the strategic environment, I calculate how many performance links a mental model of M strategic choices will consider. I find that across all strategic environments, given a mental model of size M, theories will be simpler representations than the equivalent associative mental model.

In developing the intuition for why cause-and-effect performance links will more simply represent the strategic environment, I return to our motivating example. A startup founder is using a mental model to select the features of her AI model for cancer detection: data type, learning type and evaluation metric. In her mental model, she considers the performance links between two features, the data type and the learning type (M=2), selecting the evaluation metric independently.

Using an associative mental model to choose the data and learning type (with M = 2), she considers performance links between the features. Because the true strategic environment has causal performance links between choices, the founder representing an association between these two strategic choices is representing two possible performance links between them. Either the data type is a performance antecedent of the learning type selection, or vice versa, or both. Thus, even though an association between two features may seem like one connection, by not specifying which feature affects the performance of the other, the founder is implicitly representing two possible causal performance links in her associative mental model.

In general, associative models include all possible performance links between choices since they specifically suggest a performance relationship between choices, but don't specify a causal direction to the performance dependency between them. So, for an associative model with M choices, the total number of performance links follows the formula shown in equation 1.

$$I_{associative} \; = \; M \; * \; (M - 1) \tag{1}$$

When using a theory to choose the features for her AI model (M=2), the startup founder instead considers how the choice of the data type impacts the performance of the learning type selection. Since this is a directed link between the data and learning type, the decision-maker is only considering one proposed performance link between these two features, that of the data type's impact on the performance of the learning type.

In general, as longer chains of cause-and-effect links are represented, the decision-maker is implicitly considering the performance consequences of all strategic choices that come prior to each strategic choice in the mental model. Thus, the number of performance links considered by a mental model with cause-and-effect performance links follows the formula of equation 2 below.

$$I_{cause-and-effect} = \sum_{j=1}^{M} j - 1 = \frac{M*(M-1)}{2} \tag{2}$$

Since equations 1 and 2 show the number of performance links in associative and cause-and-effect mental models based on the number of choices (M) included in the model, it is possible to compare them to calculate when causal models are simpler. Specifically, equation 3 substitutes the number of links in associative models into equation 2 and shows for mental models of a given number of choices (M), causal models contain half as many links as associative ones.

$$I_{cause-and-effect} = \frac{I_{associative}}{2}, \text{ thus } I_{cause-and-effect} < I_{associative} \text{ } for \text{ } a \text{ } given \text{ } M \tag{3}$$

Figure 5 visualizes equation 1 and equation 2 across values of M, showing that the number of performance links represented by cause-and-effect mental models ($I_{cause-and-effect}$) will always be half that of the number of performance links of the equivalent associative mental models ($I_{associative}$).

FIGURE 5

Simplicity of Representation by Size (M) and Link Type
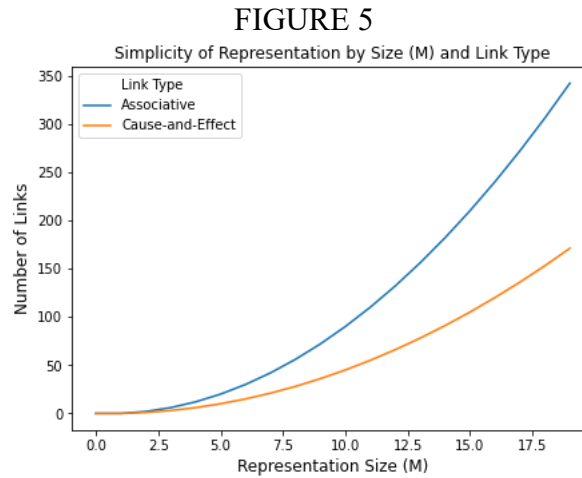
Because the number of performance links represented in each mental model does not depend on the number of strategic choices (N) or the number of actual performance links between these choices (K), this finding will hold across all strategic environments. Thus, for a given set of M strategic choices, cause-and-effect models will always be simpler representations

15

than associative models. However, while the simplicity of a representation does not depend on the environment's complexity, the accuracy of each model does — a result which I turn to next.

3.3.2 The Effect of a Mental Model's Performance Link Type on the Accuracy of the Representation

To model how accurately a mental model represents the strategic environment, I measure the degree to which its performance links match the true performance links in the environment. Since the environment is based on an NK-landscape—an NP-hard problem with no generalizable solution—I use both a mathematical 'in expectation' calculation and a computer simulation to estimate accuracy. I find that which type of model is more accurate depends on the environment's complexity. When there are more strategic choices (higher N), theories are more accurate. But when there are more links between choices (higher K), associative models are more accurate.

First, I calculate how accurate mental models with cause-and-effect and associative performance links are on average. The accuracy of a mental model's performance links is, just like in accuracy calculations in prediction tasks, the number of correct performance links represented (both links that are present and absent), divided by the number of total performance links in the representation (i.e. Baldi et al., 2000; equation 4a). For any given environment, the number of total performance links in the representation is every possible link between the N strategic choices and the N-1 other strategic choices in the environment (equation 4b). The actual number of performance links that exist in the environment is the number of interdependencies K, times the number of strategic choices N (equation 4c).

In our example, there are 3 total choices in the environment (N=3), which makes the total number of possible performance links in the mental model six (equation 4b). Because all three of

the makeup products (N=3) depend on every other product's chemical base selection (K=2), every possible performance link between the products in the environment exists (K*N also equals six, equation 4c). Because the number of actual performance links (equation 4c) equals the number of total performance links (equation 4b), every performance link included in the mental model (I), will link increase its accuracy by 1 divided by N*(N-1)—the total number of possible links in the environment.

This relationship between the number of links in the mental model and the accuracy of the mental model can be generalized using the example's intuition. To find the expected number of correct links in a mental model, multiply the chance that any given performance link will be correct, which is the number of actual performance links divided by the number of possible performance links, by the total number of links in the model (I), as shown in equation 4d. To find the expected number of correct performance non-links modeled, subtract the non-links that were modeled as links from the actual number of non-links in the strategic environment, as shown in equation 4e.

Equation 4: Accuracy of Performance Links Represented in a Mental Model with I

Performance Links

$$\text{Accuracy of performance links}, Acc = \frac{TP+TN}{TP+FP+TN+FN} \tag{4a}$$

$$\text{Total possible number of performance links}, TPT = TP + FP + TN + FN = N(N-1) \tag{4b}$$

$$\text{Total actual number of performance links}, TT = KN \tag{4c}$$

$$\text{Expected number of correct performance links}, TP = I * \frac{KN}{N(N-1)} \tag{4d}$$

$$\text{Expected number of correct non-links}, TN = N(N-1) - KN - (I - (\frac{I*K}{N-1})) \tag{4e}$$

Overall, equation 4 calculates the expected accuracy of a mental model in terms of characteristics of the strategic environment (N and K) and the number of performance links in

the mental model (I). Since equations 1 and 2 give the number of performance links (I) for associative and cause-and-effect models for a given M, I can plug these values into equation 4 to find when cause-and-effect models are more accurate. This result is calculated in equation 5 (see Appendix B for all steps of the solution), an inequality that shows that for environments where ½ is greater than K/(N-1), cause-and-effect models are more accurate than associative ones.

$$\frac{1}{2} > \frac{K}{N-1} \tag{5}$$

Second, because the strategic environment is assumed to be an NK-landscape, a type of NP-hard problem with no general solution, I use computer simulations to compare how accurately a mental model of size M—one with associative performance links and one with a *random set* of cause-and-effect performance links—represent a specific NK-landscape. Figure 6 shows these results: Figure 6a compares model link accuracy as the number of strategic choices (N) increases (with K = 3), and Figure 6b compares link accuracy as the number of performance links (K) increases (with N = 10).

FIGURE 6
FIGURE 6A: Accuracy of Links by Representation Size (M) and Number of Strategic Choices (N) for Mental Models of Varying Link Types
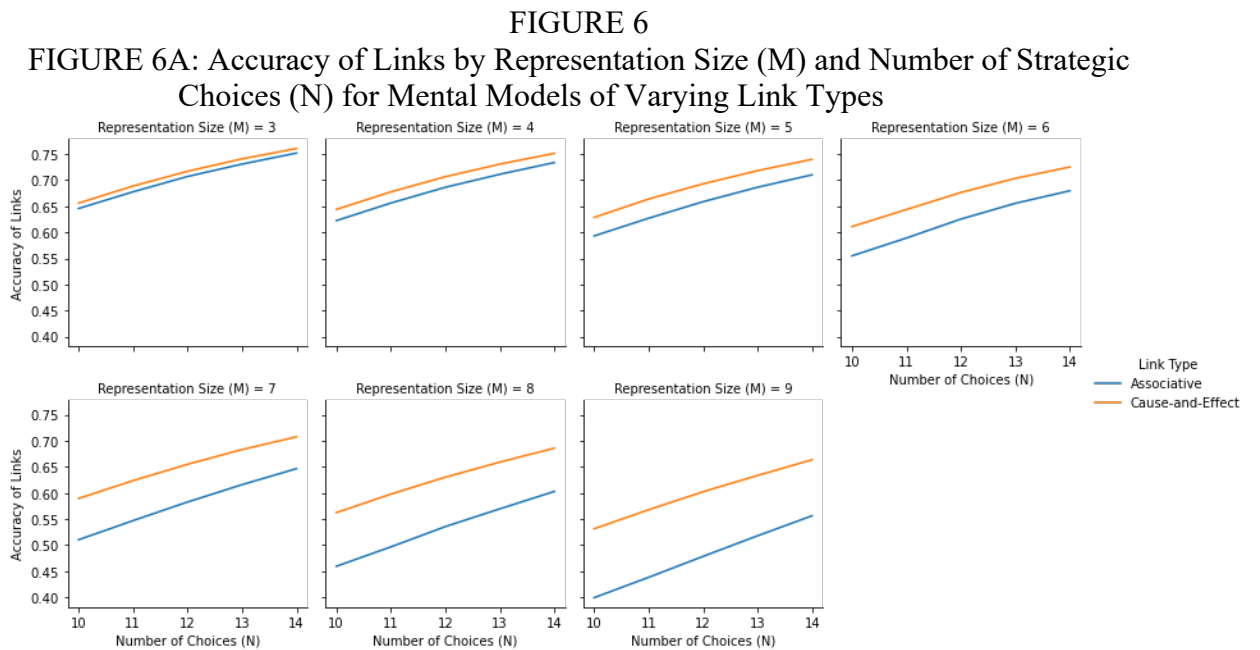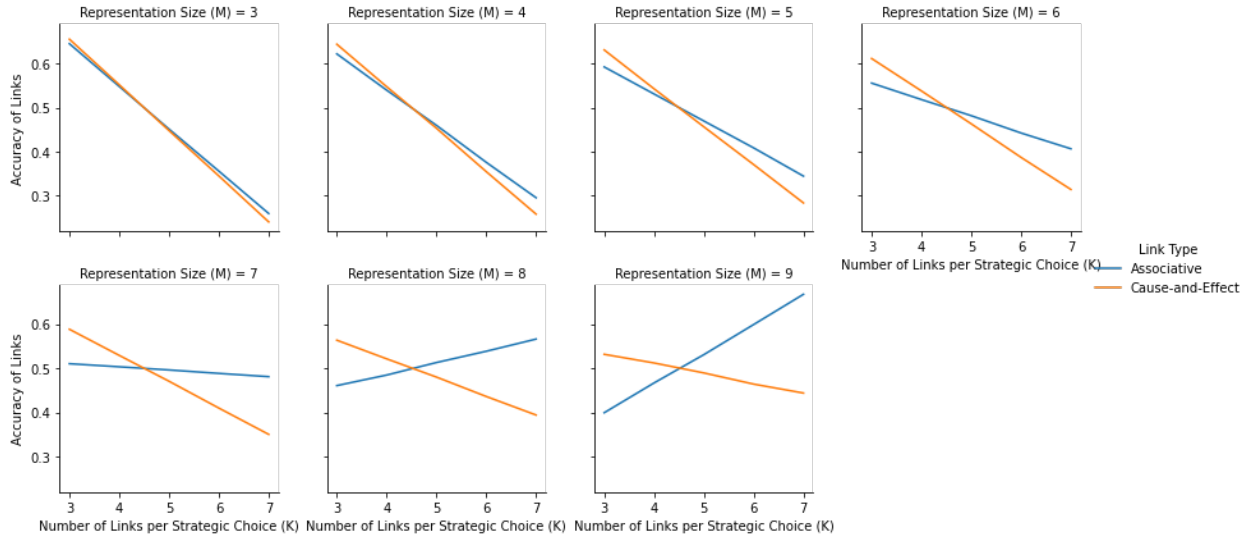

18

FIGURE 6B: Accuracy of Links by Representation Size (M) and Number of Performance Links Between Strategic Choices (K) for Mental Models of Varying Link Types

Supporting the calculations of equations 4 and 5, the computer simulations of Figure 6a above show that as the number of strategic choices (N) increases, cause-and-effect models are more accurate across all mental model sizes (M). When the number of links between choices (K) increases (with N fixed at 10), equation 5 predicts that cause-and-effect models are more accurate when K is below 5 and less accurate when K is above 5—which is exactly what Figure 6b shows.

Now that I've compared how associative mental models and theories differ in simply and accurately representing the strategic environment, I turn to how these differences affect performance outcomes.

3.4 The Performance Consequences of the Parsimonious Representations by Performance Link Type of Mental Models in Environments of Increasing Complexity

To model the performance consequences of using an associative vs. causal mental model to select a strategy in a complex strategic environment, I run 1000 iterations for 100 times periods of mental models of varying size M (from 3 to 9), where both associative mental models and theories search for strategies across NK-landscapes as described above.

Figure 7 shows the performance of strategies found by decision-makers using associative vs. causal mental model of size (M). (with N = 10 and K = 3). These results are stable across strategic environments of varying N and K, as shown in Appendix C.
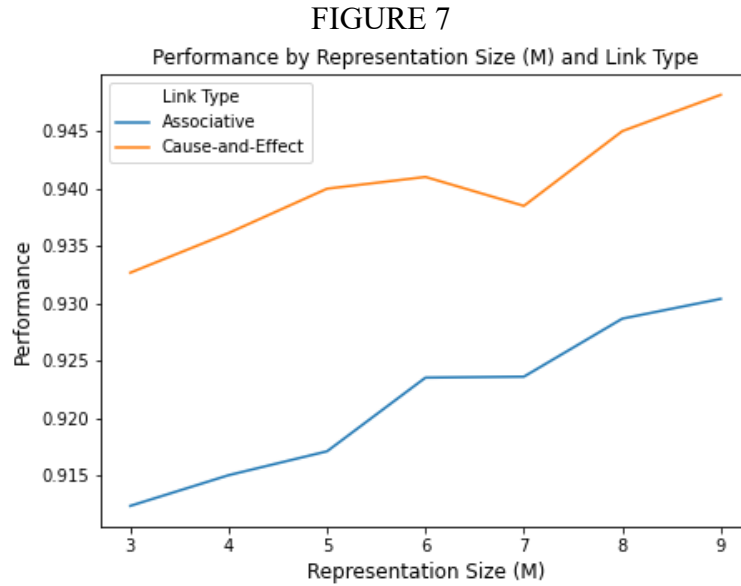
FIGURE 7

Figure 7 shows two important results. First, mental models that contain more strategic choices (high M) perform better than those that contain fewer (low M). This is line with previous research that finds when there are dimensions in the strategic environment do not vary in their importance for determining performance, maximizing the number of strategic choices in the representation is beneficial (Csaszar & Levinthal, 2016). However, this does not mean that simplicity does not have a role in improving performance, because the second important observation from Figure 7 is that for mental models at any given size M, theories outperform associative mental models.

In the below sections, I explore how the trade-off that theories make to represent fewer performance links for a given representation of size M when compared to associative mental models leads them to find better strategies across complex strategic environments.

3.4.1 Theories Represent Fewer Performance Links at Every Sized Representation, Making Them

More Performant

I propose that theories will outperform associative mental models because they represent

fewer performance relationships between strategic choices, allowing decision-makers to more

effectively use performance feedback to search for better strategies. The above result shows that

theories do outperform associative mental models across all strategic environments. However, to

explore whether the mechanism of simplicity of representation is a consistent explanation of this

result, I show how the number of links in the mental model (I) is related to the performance of

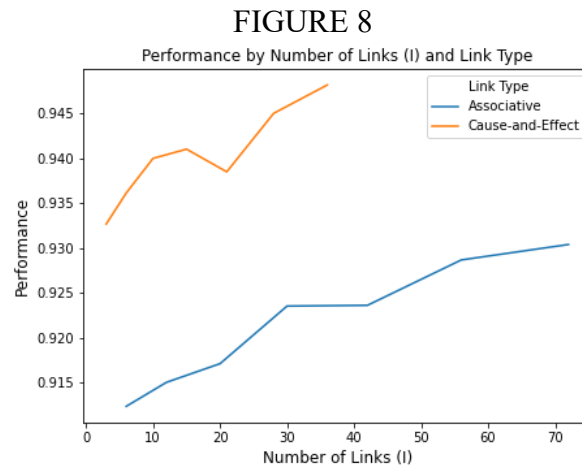the strategies found by each mental model type.

FIGURE 8

Figure 8 shows the relationship between the performance of strategies found and the

number of links in a mental model (I) by type of representation (associative vs. causal) for a

strategic environment with N = 10 and K = 3. There are two important observations from this

graph. First, as the number of links represented in the mental model increases, the performance

of strategies found increases. Second, for any given number of links represented in a mental

model (I) a theory outperforms an associative model. While these results may appear to suggest

that theories are more performant when representing any given number of performance links,

because the number of performance links represented in a mental model (I) is a function of the

number of strategic choices in the model (M), these results strongly support simplicity of representation as an explanation for performance.

FIGURE 9

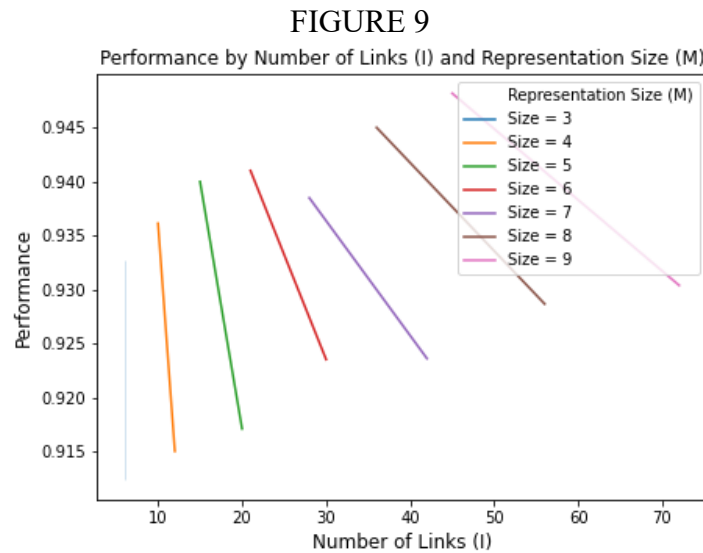Performance by Number of Links (I) and Representation Size (M)

Figure 9 shows why this is the case, by graphing the performance of strategies selected and the number of performance links in a representation (I) by the number of strategic choices in the mental model (M). Once again, there are two important observations to make about this figure. First, for a mental model of a given size (M), it is always more performant to have fewer number of links represented (I), which is shown by the negative slope of the performance line for each mental model size (M). Second, tying the results of Figure 9 to Figure 8 and the calculations of equations 1 and 2, theories represent fewer performance links (I) than associative mental models at a given mental model size (M). Thus, the simplicity of representation of performance links by theories is a plausible explanation for the performance differences seen between the strategies selected by theories and associative mental models.

Because the number of links in a representation isn't dependent on features of the environment (N or K), these results hold across all strategic environments as shown in Appendix D. Thus, theories are more simple representations of strategic environments, and this simplicity finds more successful strategies across environments of increasing complexity.

22

3.4.2 When A Trade-Off Between Simplicity and Accuracy Arise, Theories Make More

Performant Choice of Simplicity

While simplicity appears to be a consistent explanation of the performance difference between strategies found by theories and associative mental models, previous work has shown that the other major predictor of mental models' success in selecting performant strategies is how accurately these models represent the strategic environment (Gary & Wood, 2010; Csaszar & Levinthal, 2016). Indeed, across a wide range of strategic environments, the above theoretical development suggests that theories will be more accurate representations of the strategic environment than associative mental models. Using mental models searching for strategies as described above, I calculate the actual accuracy of performance links in associative and causal mental models representing NK-landscape strategic environments and the performance of the strategies selected by these models.
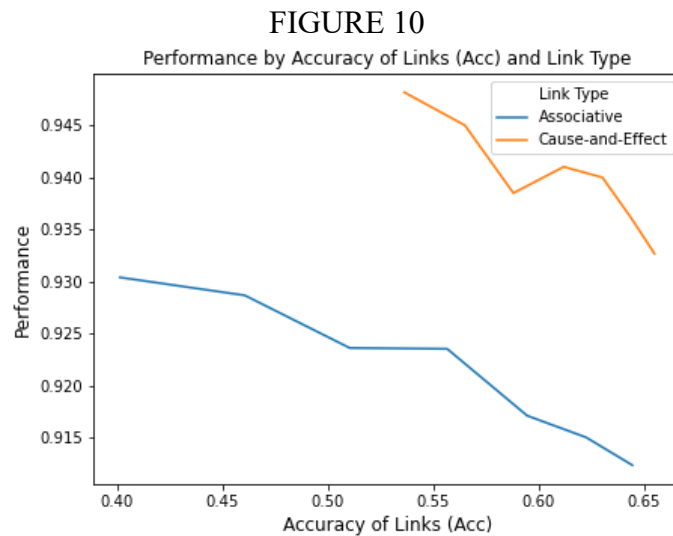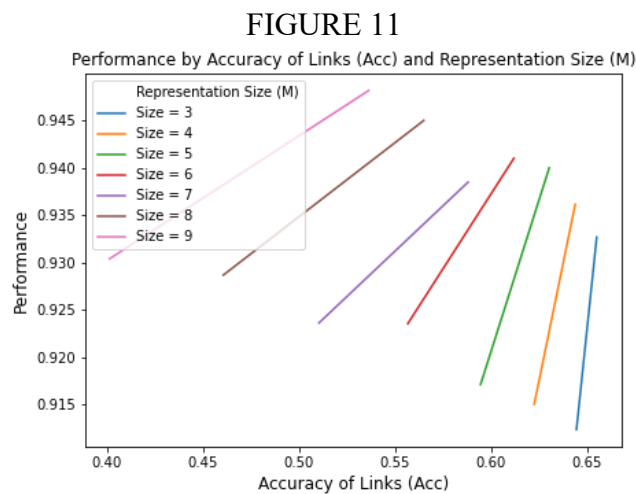
FIGURE 10

Figure 10 shows the relationship between performance and accuracy of representation (Acc) by type of mental model (associative vs. theory) for a strategic environment of N = 10 and K = 3. Like the results with simplicity in Figure 8, there are two important observations to make with Figure 10. First, as the accuracy in the performance links represented in the mental model

increases, the performance of strategies found appears to decrease. Second, for any given

accuracy of links represented in a mental model (Acc), a theory outperforms an associative

mental model. These results could support accuracy of representation as an explanation for

performance differences between theories and associative mental models, but they appear to

somewhat counterintuitively suggest that accuracy in representing the strategic environment does

not improve the performance of strategies found.

Exemplifying Simpson's paradox, Figure 11 shows why accuracy in representation is

actually associated with higher performance, by graphing the performance of strategies selected

by accuracy of representation (Acc) and by size of the mental model (M). Two observations

follow. First, for a mental model of a given size (M), it is more performant to represent the

environment more accurately (Acc), which is shown by the positive slope of the performance

line for each mental model size (M). Second, tying the results of Figure 11 to Figure 10 and the

calculations of equations 4 and 5, in this strategic environment ($N = 10$, $K = 3$) theories represent

the environment more accurately than associative mental models at every mental model size (M).

FIGURE 11

Performance by Accuracy of Links (Acc) and Representation Size (M)

However, the accuracy of a mental model depends on the strategic environment (both N

and K). Particularly, equation 5 proposes that for environments that are increasingly complex in

terms of the number of performance links they contain (high K), theories will be less accurate than associative mental models at representing the strategic environment. As predicted by equation 5, Appendix E shows that across environments of increasing number of strategic choices (N), theories remain more accurate representations of the strategic environment.

Figure 12, however, shows the relationship between performance and accuracy of representation (Acc) by type of mental model (associative vs. theory) for a strategic environment of N = 10 where K increases from 3 to 7. Consider the two following observations. First, when K is low (K < 5), theories are more accurate and more performant than associative mental models. Second, when K is high (K >= 5), theories are less accurate and more performant than associative mental models. Like above, I can show these results by size of the representation (M) to explain further.

FIGURE 12: Performance of Mental Models by Accuracy (Acc) and Link Type Across Environments of Increasing Number of Performance Links Between Strategic Choices (K)
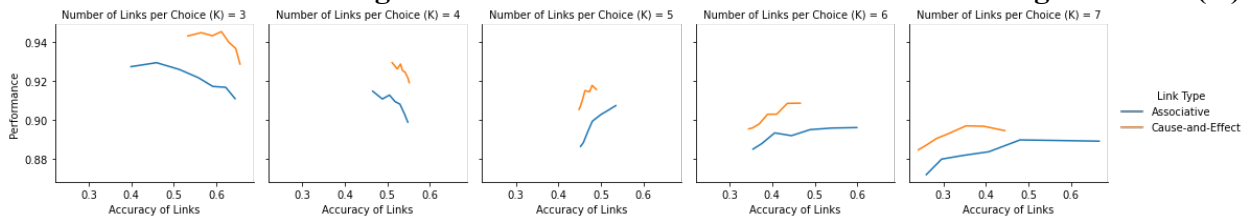

Figure 13, graphs performance and accuracy by size of representation (M) across strategic environments with N = 10 and increasing value of K (from 3 to 7). Consistent with the results above and the predictions of equation 5, observe the following two trends. First, when K is low (K<5), it is more performant to represent the environment more accurately (Acc), which is shown by the positive slope of the performance line for each mental model size (M). Second, when K is high (K>= 5), it is more performant to represent the environment less accurately (Acc), which is shown by the negative slope of the performance line for each mental model size

(M). Thus, across strategic environments of varying complexity, accuracy of representation is

not a consistent explanation for why theories outperform associative mental models.

FIGURE 13: Performance of Mental Models by Accuracy (Acc) and Size (M) Across Environments of Increasing Number of Performance Links Between Strategic Choices (K)
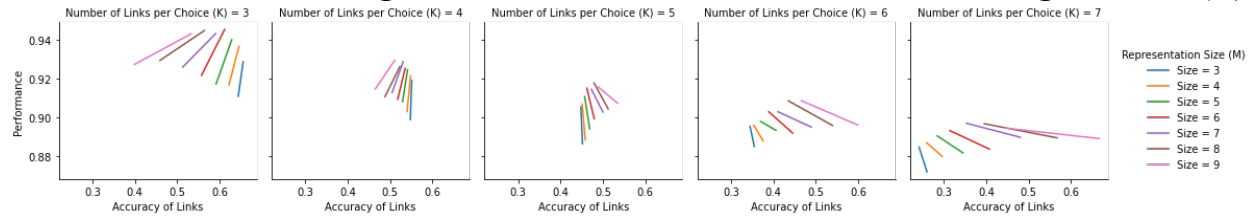
Overall, my results suggest that theories, by representing the performance links of the

strategic environment more simply, find more successful strategies than associative mental

models. These results are consistent with previous work on mental models that suggests that one

of the major benefits of these representations is simplifying environmental complexity to

facilitate effective search for strategies (Gavetti & Levinthal, 2000). However, theories simplify

the strategic environment more effectively than associative models by reducing the number of

links between choices, rather than the number of choices themselves, leading to better strategies.

4. DISCUSSION AND CONCLUSION

The cause-and-effect performance links in theories help decision-makers choose

successful strategies by being more parsimonious models of the strategic environment than

associative mental models. The simple, and often more accurate, representation of performance

links by theories helps decision-makers consider fewer conditional dependencies between

strategic choices, better using prior performance feedback to select strategies. For highly

interdependent environments where there is a trade-off between accuracy and simplicity, theories

make the more performant choice of representing the strategic environment more simply, helping

decision-makers find better strategies.

26

Overall, the above results, which suggest that theories are always better than associative models at finding successful strategies in complex strategic environments, prompt a series of questions and, hopefully, a series of paths for future research. In the sections below I ask several of these questions and suggest some promising areas for addressing them through further work.

4.1 Simplicity Over Accuracy?: Simple, Causal Mental Models Provide Structure that Helps Incorporate Performance Feedback

A key question these results raise is: why does simplicity outperform accuracy in strategic environments that require a trade-off between the two? Research on mental models in strategy, and on human cognition more generally, shows that people often use simple models to make effective decisions in complex situations (Bingham & Eisenhardt, 2011; Schwenk, 1984; Simon, 1977; Gigerenzer, 2008). However, there is a debate over whether causal reasoning actually improves decision-making (Sperber et al., 1995; Mercier & Sperber, 2017).

My findings suggest that cause-and-effect links make mental models simpler in a way that helps decision-makers better use performance feedback in complex environments. In these settings, the main challenge is turning past feedback into better future strategies (Siggelkow, 2011; March, 1991). When there's no true solution to environmental complexity (strategy as an NP-hard problem), mental models with cause-and-effect performance links provide the next best thing — a simple structure of sequential choices that makes performance feedback easier to understand (see Pearl, 2009; Denrell et al., 2004).

This result can be seen as a formalization of the empirical findings in the theory-based view of strategy, where holding a theory improves experimentation and allows for the application of the scientific method to organizational strategy (Camuffo et al., 2024; Camuffo et

al., 2020). I outline two implications of the simplicity of theories outperforming more accurate associative mental models below.

First, my results suggest that evaluations of the quality of a theory should be separated from evaluations of the performance benefits of holding one. While previous work shows that iterating with theories can improve the accuracy of performance understandings (Camuffo et al., 2024; Ehrig & Schmidt, 2022), my work shows that performance gains may come from how the theory structures decision-making—which does not depend on how accurate the theory is. Future research should separate these two types of benefits.

Second, I propose that theories lead decision-makers to choose strategies in a step-by-step (sequential) way, while associative models lead to all-at-once (simultaneous) decisions. Beyond validating this basic difference, future research should explore how these mental models create distinct decision-making processes, especially in how strategic choices are ordered.

4.2 Does it Really Not Matter if My Theory is "Good"?

Because my model compares the average performance of theories versus associative mental models, it doesn't show how individual theories—each with different levels of accuracy—perform. This raises another question: does a theory's performance depend on how accurate, or "good", it is?

While my results show that accuracy of representation doesn't always generate performance, they cannot directly answer this question because of how the baseline version of the NK-landscape works. In this model, it is assumed that every strategic choice (N) depends on K others, and while some choices have more impact than others through the weight matrix (W), it is difficult to conceptualize any strategic choice as more or less causally central to the performance of a strategy in this model. Because the strategic environment is largely causally

non-ordered, most theories which causally order strategic choices are equally as accurate or "good." Future research should explore how theories affect decision-making in more structured (or "ordered") environments (see Csaszar & Levinthal, 2016 use of parameter W for an example)—where some theories clearly reflect the environment better than others and might perform differently as a result.

4.3 What is the Role of Theories in Generating Group Consensus?

Recent research on strategic decision-making in organizations has emphasized the need for group consensus (Adner & Levinthal, 2024; Carroll & Sørensen, 2021). My findings show that theories help individuals navigate complex environments, but they may also lead to more disagreement through what I call the "chicken-and-egg effect". While everyone may agree two things are related, theories which assign a clear causal direction (e.g., chicken before egg or vice versa) are more likely to spark conflict. Future research should explore how assigning causal direction can create disagreement in groups and look for ways to reduce this downside of using theories in strategic decision-making.

4.4. Conclusion: Any Old Theory Will Do?

An old poem adopted by organizational scholars told the story of lost soldiers who found their way back to camp using a map of the wrong mountain range—and from it came the phrase "any old map will do" (Holub, 1977). Similarly, my results show that even a randomly generated theory can outperform a more accurate but complex associative model in hard-to-learn environments. This suggests that having a simple theory to guide strategy may be like having the wrong map: helpful, even though it is inaccurate.

Just as water flows towards valleys in a mountain range, complex strategic environments still have patterns linking choices to performance. A mental model that assumes a clear (even if

incorrect) path through the environment may better detect these patterns than one that tries to track everything (see Denrell et al., 2004). This helps explain why simpler theories can outperform more accurate models when the environment is too complex to fully understand—a finding supported by broader research on heuristics and cognitive shortcuts (Bingham & Eisenhardt, 2011; Simon, 1977; Gigerenzer, 2008). Overall, my model shows that theories find more successful strategies because they offer a simpler way to navigate complexity.

WORK CITED

Adner, Ron, and Daniel A. Levinthal. "Strategy Experiments in Nonexperimental Settings: Challenges of Theory, Inference, and Persuasion in Business Strategy." *Strategy Science*, Oct. 2024. *pubsonline.informs.org (Atypon)*, https://doi.org/10.1287/stsc.2024.0164.

Baldi, Pierre, et al. "Assessing the Accuracy of Prediction Algorithms for Classification: An Overview." *Bioinformatics*, vol. 16, no. 5, May 2000, pp. 412–24. *Silverchair*, https://doi.org/10.1093/bioinformatics/16.5.412.

Baumann, Oliver, et al. "Effective Search in Rugged Performance Landscapes: A Review and Outlook." *Journal of Management*, vol. 45, no. 1, Jan. 2019, pp. 285–318. *SAGE Journals*, https://doi.org/10.1177/0149206318808594.

Bingham, Christopher B., and Kathleen M. Eisenhardt. "Rational Heuristics: The 'Simple Rules' That Strategists Learn from Process Experience." *Strategic Management Journal*, vol. 32, no. 13, 2011, pp. 1437–64. *Wiley Online Library*, https://doi.org/10.1002/smj.965.

Brusoni, Stefano, et al. "The Value and Costs of Modularity: A Problem-Solving Perspective." *European Management Review*, vol. 4, no. 2, 2007, pp. 121–32. *Wiley Online Library*, https://doi.org/10.1057/palgrave.emr.1500079.

Camuffo, Arnaldo, Alessandro Cordova, et al. "A Scientific Approach to Entrepreneurial Decision Making: Evidence from a Randomized Control Trial." *Management Science*, vol. 66, no. 2, Feb. 2020, pp. 564–86. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/mnsc.2018.3249.

Camuffo, Arnaldo, Alfonso Gambardella, et al. "Theory-Driven Strategic Management Decisions." *Strategy Science*, Nov. 2024. *pubsonline.informs.org (Atypon)*, https://doi.org/10.1287/stsc.2024.0173.

Carroll, Glenn R., and Jesper B. Sørensen. *Making Great Strategy: Arguing for Organizational Advantage*. Columbia Business School Publishing, 2021.

Choi, Jaeho, and Daniel Levinthal. "Wisdom in the Wild: Generalization and Adaptive Dynamics." *Organization Science*, vol. 34, no. 3, 2023, pp. 1073–89. *APA PsycNet*, https://doi.org/10.1287/orsc.2022.1609.

Csaszar, Felipe A., and Daniel A. Levinthal. "Mental Representation and the Discovery of New Strategies." *Strategic Management Journal*, vol. 37, no. 10, 2016, pp. 2031–49. *Wiley Online Library*, https://doi.org/10.1002/smj.2440.

Csaszar, Felipe A., and James Ostler. "A Contingency Theory of Representational Complexity in Organizations." *Organization Science*, vol. 31, no. 5, Sep. 2020, pp. 1198–219. *pubsonline.informs.org (Atypon)*, https://doi.org/10.1287/orsc.2019.1346.

Denrell, Jerker., et al. "From T-Mazes to Labyrinths: Learning from Model-Based Feedback." *Management Science*, vol. 50, no. 10, Oct. 2004, pp. 1366–78. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/mnsc.1040.0271.

Ehrig, Timo, and Jens Schmidt. "Theory-Based Learning and Experimentation: How Strategists Can Systematically Generate Knowledge at the Edge between the Known and the Unknown." *Strategic Management Journal*, vol. 43, no. 7, 2022, pp. 1287–318. *Wiley Online Library*, https://doi.org/10.1002/smj.3381.

Felin, Teppo, et al. "Theory-Based Decisions: Foundations and Introduction." *Strategy Science*, vol. 9, no. 4, Dec. 2024, pp. 297–310. *pubsonline.informs.org (Atypon)*, https://doi.org/10.1287/stsc.2024.intro.v9.n4.

Ganco, Martin, and Glenn Hoetker. "NK Modeling Methodology in the Strategy Literature: Bounded Search on a Rugged Landscape." *Research Methodology in Strategy and Management*, vol. 5, Emerald Group Publishing Limited, 2009, pp. 237–68. world, *www-emerald-com.stanford.idm.oclc.org*, https://www.emerald.com/insight/content/doi/10.1108/s1479-8387(2009)0000005010/full/html.

Gary, Michael Shayne, and Robert E. Wood. "Mental Models, Decision Rules, and Performance Heterogeneity." *Strategic Management Journal*, vol. 32, no. 6, 2011, pp. 569–94. *Wiley Online Library*, https://doi.org/10.1002/smj.899.

Gavetti, Giovanni, and Daniel Levinthal. "Looking Forward and Looking Backward: Cognitive and Experiential Search." *Administrative Science Quarterly*, vol. 45, no. 1, Mar. 2000, pp. 113–37. *SAGE Journals*, https://doi.org/10.2307/2666981.

Gavetti, Giovanni, and Jan W. Rivkin. "On the Origin of Strategy: Action and Cognition over Time." *Organization Science*, Jun. 2007. world, *pubsonline.informs.org*, https://doi.org/10.1287/orsc.1070.0282.

Georgakopoulos, Diimitrios, et al. "An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure." *Distributed and Parallel Databases*, vol. 3, no. 2, Apr. 1995, pp. 119–53. *Springer Link*, https://doi.org/10.1007/BF01277643.

Gigerenzer, Gerd. "Why Heuristics Work." *Perspectives on Psychological Science*, vol. 3, no. 1, Jan. 2008, pp. 20–29. *SAGE Journals*, https://doi.org/10.1111/j.1745-6916.2008.00058.x.

Hochba, Dorit S. "Approximation Algorithms for NP-Hard Problems." *SIGACT News*, vol. 28, no. 2, Jun. 1997, pp. 40–52. *ACM Digital Library*, https://doi.org/10.1145/261342.571216.

Holub, Miroslav. "Brief Thoughts on Maps." *A Geological Miscellany*, Princeton University Press, 2014 (written 1977), pp. 3–4. *www.degruyterbrill.com*, https://www.degruyterbrill.com/document/doi/10.1515/9781400857913.3/html?lang=en&srsltid=AfmBOoqGGqERqKdmP9dycYwPeIE12LRL-Qzzlp8SDTpph1dAuu2P2Ts2.

Kaplan, Sarah. "Research in Cognition and Strategy: Reflections on Two Decades of Progress and a Look to the Future." *Journal of Management Studies*, vol. 48, no. 3, 2011, pp. 665–95. *Wiley Online Library*, https://doi.org/10.1111/j.1467-6486.2010.00983.x.

Levinthal, Daniel A. "Adaptation on Rugged Landscapes." *Management Science*, vol. 43, no. 7, Jul. 1997, pp. 934–50. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/mnsc.43.7.934.

March, James G. "Exploration and Exploitation in Organizational Learning." *Organization Science*, vol. 2, no. 1, Feb. 1991, pp. 71–87. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/orsc.2.1.71.

Martignoni, Dirk, et al. "Consequences of Misspecified Mental Models: Contrasting Effects and the Role of Cognitive Fit." *Strategic Management Journal*, vol. 37, no. 13, 2016, pp. 2545–68. *Wiley Online Library*, https://doi.org/10.1002/smj.2479.

Mercier, Hugo, and Dan Sperber. *The Enigma of Reason*. Harvard University Press, 2017, pp. vi, 396. *APA PsycNet*, https://doi.org/10.4159/9780674977860.

Nelson, Richard R. *An Evolutionary Theory of Economic Change*. Belknap Press, 2004.

Pearl, Judea. *Causality*. 2nd ed., Cambridge University Press, 2009. *Cambridge University Press*, https://doi.org/10.1017/CBO9780511803161.

Porac, Joseph F., et al. "Competitive Groups as Cognitive Communities: The Case of Scottish Knitwear Manufacturers." *Journal of Management Studies*, vol. 26, no. 4, 1989, pp. 397–416. *Wiley Online Library*, https://doi.org/10.1111/j.1467-6486.1989.tb00736.x.

Rivkin, Jan W. "Imitation of Complex Strategies." *Management Science*, vol. 46, no. 6, Jun. 2000, pp. 824–44. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/mnsc.46.6.824.11940.

Schwenk, Charles R. "Cognitive Simplification Processes in Strategic Decision-Making." *Strategic Management Journal*, vol. 5, no. 2, 1984, pp. 111–28. *Wiley Online Library*, https://doi.org/10.1002/smj.4250050203.

Siggelkow, Nicolaj. "Firms as Systems of Interdependent Choices." *Journal of Management Studies*, vol. 48, no. 5, 2011, pp. 1126–40. *Wiley Online Library*, https://doi.org/10.1111/j.1467-6486.2011.01010.x.

Simon, Herbert A. "The Logic of Heuristic Decision Making." *Models of Discovery*, Springer, Dordrecht, 1977, pp. 154–75. *link-springer-com.stanford.idm.oclc.org*, https://doi.org/10.1007/978-94-010-9521-1_10.

Sorenson, Olav. "Theory, Search, and Learning." *Strategy Science*, vol. 9, no. 4, Dec. 2024, pp. 372–81. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/stsc.2024.0179.

Sperber, Dan et al. *Causal Cognition: A Multidisciplinary Debate*. Clarendon Press/Oxford University Press, 1995, pp. xx, 670.

Sperber, Dan et al. "Organizational Capabilities in Complex Worlds." *The Nature and Dynamics of Organizational Capabilities*, edited by Giovanni Dosi et al., Oxford University Press, 2001, p. 0. *Silverchair*, https://doi.org/10.1093/0199248540.003.0014.

Tripsas, Mary, and Giovanni Gavetti. "Capabilities, Cognition, and Inertia: Evidence from Digital Imaging." *Strategic Management Journal*, vol. 21, no. 10/11, 2000, pp. 1147–61.

Valentine, Jacob, et al. "The Theory-Based View and Strategic Pivots: The Effects of Theorization and Experimentation on the Type and Nature of Pivots." *Strategy Science*, vol. 9, no. 4, Dec. 2024, pp. 433–60. *pubsonline-informs-org.stanford.idm.oclc.org (Atypon)*, https://doi.org/10.1287/stsc.2024.0183.

APPENDIX

Code for this model is publicly available here: https://github.com/mhsingell/theory

Pseudo Code:
Using parameters as defined in Table 1, run for NS number of simulations.
1. Initialize NK-landscape with N and K.
2. At time T = 0, randomly select M strategic choices for mental model, an order (OR) for theory (A→B→C…→M) and an initial random strategy.
3. For each subsequent T, use the mental model or theory to select a new strategy:
    1. Start with the best prior performing strategy.
    2. Select a random strategic choice A to change from best.
    3. For mental model runs calculate the best average performing strategy conditional on the value of choice A selected in 2. for all M choices.
    4. For theory runs calculate the best average performing strategy for performance consequents of A only conditional on the value of choice A and its antecedents selected in 2. using order (OR).
    5. For all N-M strategic choices, select best choice based on independently considered performance.
4. If strategy selected in 3. has already been tried, select a random strategic choice in N to change.
5. Receive performance feedback, repeat for all T.