Bridging the gulf between executive characteristics and organizational outcomes: The intertwined roles of cognition, aggregation structure, and the environment\textsuperscript{1}

Felipe A. Csaszar
Ross School of Business, U. of Michigan
Ann Arbor, MI 48109
fcsaszar@umich.edu

October 23, 2014

\textsuperscript{1}For helpful comments, the author thanks Nicolaj Siggelkow and JP Eggers. The author also thanks seminar participants at the University of Michigan, the Organization Science Winter Conference, the Carnegie School of Organizational Learning at Asilomar, the Theoretical Organizational Models Society Meeting at New York, and the Duke Strategy Conference. The author thanks Sougata Chaudhuri and Chamithrie Wijayasena for excellent research assistance. Errors remain the author’s own.
Abstract

This paper presents a theory about how firm performance depends on cognition, aggregation structure, and the environment. The theory is developed as a formal model, and it characterizes: cognition in terms of managers’ experience and mental representations (whether they apprehend the environment from a specialist’s versus a generalist’s viewpoint); aggregation structure as decision rules (whether the firm’s decisions are made by one manager or by several managers); and the environment as the projects—of varying complexity and uncertainty—considered by the firm. The results show that, in environments of low uncertainty or high complexity, an inexperienced generalist outperforms groups of experienced specialists at managing firms; yet when the environment is uncertain but not complex, groups of experienced specialists do better. The model is used to identify industry conditions under which it may be advantageous to use co-CEOs, inexperienced CEOs, or specialist CEOs. The paper also discusses whether management education should focus on training specialists or generalists and whether firms should follow simple or complex rules.

Keywords: cognition, aggregation, mental representation, organizational structure
1 Introduction

In a recent survey of the literature on top management teams, Finkelstein et al. (2009:115) note that, despite extensive research efforts in this area, “the gulf between executive characteristics and organizational outcomes is huge.” The implication is that, although studies linking executive characteristics to organizational outcomes have found statistical regularities, the causal connections between such characteristics and outcomes are not well understood. In other words, the processes that mediate between managers’ micro characteristics and organizations’ macro outcomes are unclear.

In attempting to bridge this gulf, it is important to recognize the three main processes must mediate the effect of managers’ characteristics on organizational outcomes. As illustrated by Figure 1, managers’ characteristics and organizational outcomes are just the extremes of a larger system that includes three processes: (I) cognition, (II) aggregation, and (III) the environment. Cognition determines how managers’ characteristics (e.g., mental representation and experience) affect their opinions. Aggregation determines how managers’ opinions are combined to produce an organization-level decision (e.g., some firms let a powerful CEO dictate decisions; other firms use more egalitarian decision rules that accommodate the opinions of several managers). Finally, the environment determines how the organization’s decision are transformed into organizational outcomes (e.g., in an uncertain environment, similar decisions may lead to different outcomes; in complex environments, outcomes are highly sensitive to interactions among decision characteristics).

![Figure 1: Processes connecting executive characteristics to organizational outcomes.](image)

Processes I and III have been thoroughly examined in the fields of psychology and economics, respectively. However, process II—how an organization aggregates managers’ opinions—has been much less studied (Gavetti et al. 2007, Csaszar and Eggers 2013). The insufficient attention paid to the aggregation process accounts for the present gulf between executive characteristics and organizational outcomes. In particular, because aggregation occurs between processes I
and III, failing to understand it prevents one from reliably predicting the effect of upstream changes on downstream results (see Figure 1). So in order to bridge the gulf between executive characteristics and organizational outcomes, it is necessary to develop a theory about processes I, II, and III and their interactions. Such a theory is desirable not only because it provides a coherent framework for understanding how effects at different levels of organization influence organizational outcomes but also because it bears implications for practice.

A question that illustrates the practical value of an integrated theory is this: Under what conditions is it advisable to use co-CEOs, inexperienced CEOs, or specialist CEOs? A priori, common wisdom indicates that firms are better managed by one CEO (rather than two CEOs), that experience is desirable in a chief executive officer, and that it is better for the CEO to be a generalist than a specialist. However, this common wisdom has recently been challenged. First, some successful companies (Deutsche Bank, Whole Foods, Chipotle) attribute their success to having co-CEOs (Feloni 2014). Second, a common view in the Internet industry is that CEO experience is mostly detrimental: “CEOs are like professional athletes, which peak at around 25 years old” (Arrington 2011). Finally, a top US auto industry executive attributes that industry’s declining performance to the hiring of CEOs who were generalist MBAs rather than specialist engineers (Lutz 2011).

Addressing these possibilities is challenging because it requires “chaining” effects across processes I, II, and III. For instance, the effect on performance of using co-CEOs should depend on their characteristics: whether the co-CEOs have complementary knowledge (an aspect of process I), how they collaborate to reach organization-level decisions (process II), and the type of environment faced by the firm (process III). Similarly, CEO experience and specialization (which are both characteristics of process I) will affect performance differently depending on who else’s opinions are taken into account (process II) and also on the environment (process III). No framework has been developed within the extant organizations literature that could be used to analyze questions such as these, whose answers depend critically on interactions among cognition, aggregation, and the environment.

The aim of this paper is to develop a process theory (Mohr 1982) that bridges the gulf between executive characteristics and organizational outcomes. In other words, a theory that—by
describing the cause–effect relationships between the extremes of Figure 1—can explain how individual characteristics lead to organizational-level outcomes. This paper develops such a theory by complementing existing models of aggregation (which cover process II) with Brunswik’s (1952) lens model (which covers processes I and III).

The theory presented here is described by a formal mathematical model that predicts how organizational performance will be affected by the characteristics of processes I, II, and III. In particular, these characteristics are: (I) the managers’ experience and mental representation (i.e., whether they are specialists or generalists); (II) the organization’s aggregation structure—whether decisions are made by a lone manager (who might be a specialist or a generalist) or by averaging the opinions of multiple specialists; and (III) the environment’s complexity and uncertainty. The model develops a stylized yet behaviorally plausible view of organizational decision making in a manner similar to recent work in the Carnegie tradition (e.g., Knudsen and Levinthal 2007, Csaszar and Eggers 2013). In order to keep the analysis tractable, the model abstracts from incentive problems (i.e., managers are assumed to report their true opinions) and studies the simplest case of two decision makers whose task is to screen projects.

This paper contributes to the organizations literature in several ways. First, it provides a theory for studying how cognition, aggregation, and the environment jointly affect firm performance. That theory combines ideas from the Carnegie tradition and Brunswik’s lens model, thereby providing a statistical learning perspective from which to analyze problems of organization theory. Second, the paper provides testable hypotheses regarding how characteristics of the three processes affect firm performance. Namely, it predicts that an inexperienced generalist will outperform groups of experienced specialists in environments with low uncertainty or high complexity but that, when the environment is uncertain yet not complex, groups of experienced specialists are better at managing firms. Third, this work addresses the practical issues already mentioned (i.e., the conditions under which it may be advantageous to use co-CEOs, inexperienced CEOs, or specialist CEOs) as well as whether management education should focus on training specialists or generalists and whether firms should follow simple or complex rules.
2 Theoretical Motivation

The theory developed in this paper builds on the literatures addressing information aggregation in organizations and the Brunswik lens model. This section offers a brief review of both literatures, and it elaborates on how they complement each other and thus enable us to study processes I, II, and III in an integrated manner.

2.1 Information Aggregation in Organizations

The literature on information aggregation in organizations (hereafter simply “aggregation”) studies the mechanisms by which organizations combine the opinions of multiple members into an organization-level decision (i.e., process II in Figure 1). Examples of aggregation include a top management team deciding which product to launch next, the partners of a venture capital firm deciding whether or not to invest in a particular start-up, and a board of directors deciding which candidate to hire for the CEO position. In each case, the organization aggregates the (possibly conflicting) opinions of several individuals into a group-level decision.

Organizations aggregate the opinions of multiple individuals in the hope that the group’s decision will be better than the decision that one individual alone could make. Studying aggregation is a challenge because there are innumerable methods of aggregating information (e.g., different ways of voting, delegating, deliberating, dividing a problem among individuals, etc.) and because the performance of these methods depends on characteristics of the decision makers and the task environment. Studying aggregation is relevant because it addresses a phenomenon that is pervasive and consequential: all organizations aggregate information, and different ways of aggregating information can lead to vastly different outcomes.

The research on aggregation is strongly rooted in the Carnegie tradition. In the foundational work of this tradition, Simon notes that even though individuals have limited cognitive capabilities, organizations can overcome these limits by designing appropriate decision-making structures (Simon 1947/1997:92–93). Seminal studies on aggregation that stem from this tradition are those by Cohen et al. (1972), who study aggregation under changing membership and conflicting preferences, and March (1991), who studies how varying degrees of mutual learning
in the aggregation process affect an organization’s ability to explore new possibilities. Recent research along these lines has examined the effect of varying specific characteristics of the aggregation process—such as centralization and decentralization (Rivkin and Siggelkow 2003), individuals’ connectivity (Fang et al. 2010), and the number of performance dimensions along which organization-level decisions are assessed (Ethiraj and Levinthal 2009).

Recent research on organizations has made some headway on solving aggregation problems by analyzing such stylized aggregation structures as voting, averaging of opinions, and delegation. This research has extended and adapted to organizational settings ideas from the economics literature on voting (Sah and Stiglitz 1986, 1988) and from the group decision-making literature on averaging and delegation (see, e.g., Kerr and Tindale 2004, Hastie and Kameda 2005). Examples of this type of work include examining how aggregation structure affects the ability of firms to explore and exploit (Knudsen and Levinthal 2007, Csaszar 2013); showing how to create arbitrarily reliable aggregation structures out of unreliable individuals (Christensen and Knudsen 2010); and studying how expertise heterogeneity and environmental change affect the relative merits of voting, averaging, and delegation (Csaszar and Eggers 2013).

An important assumption underlying most research on aggregation is that individuals are fallible; that is, individuals’ opinions about what is best for the organization may be wrong because individuals are neither perfectly rational nor perfectly informed. The fallibility assumption is typically formulated as some randomness in the process of forming opinions. Thus models of voting typically assume that with probability $\alpha$ an individual will not vote for the best alternative. Similarly, models of averaging often assume that, when assessing an alternative of “real” quality $q$, individuals will perceive a quality of $q \pm \sigma$. However, these models do not identify sources of these error terms ($\alpha$ and $\sigma$)—how cognitive characteristics affect individuals’ fallibility is beyond the scope of these models.¹

In short, research on aggregation covers one part of the gulf between individual characteristics and organizational outcomes—namely, how opinions of multiple individuals are turned into

¹Organizational economics has also studied models of aggregation (for a survey, see Garicano and Van Zandt 2013). The main focus of this literature has been to characterize optimal ways of decomposing knowledge-intensive tasks in hierarchies. As the models in the Carnegie tradition, this literature has also assumed fallibility and aligned incentives. An interesting but largely unexplored line of research is studying aggregation when individuals exhibit both fallibility and misaligned incentives (for a notable exception, see Kretschmer and Puranam 2008).
organization-level decisions. But that research does not cross the entire gulf because it fails to explain how individual opinions depend on individual characteristics (i.e., process I). Here is where the lens model of Brunswik (1952) proves valuable.

### 2.2 Brunswik’s Lens Model

Brunswik (1952:19–21) proposed his lens model as a way of explaining the process by which individuals make judgments based on multiple characteristics or *cues*. The task captured by Brunswik’s model is predicting a value when presented with some of these cues. That model conceptualizes the individual and the environment in a symmetric fashion: the individual is represented by a function that connects cues to a *predicted* value, and the environment is represented by a function that connects cues to a *real* value. For example, a manager may predict this year’s sales based on such cues as the past year’s sales ($x_1$) and growth of the economy ($x_2$) according to $y_{\text{predicted}} = x_1 + 0.5x_2$. In reality, however, the weights may differ because another cue—say, inflation ($x_3$)—also matters; in this case, the actual sales might be given by $y_{\text{real}} = x_1 + 0.3x_2 - 0.1x_3$.

Brunswik originally presented the lens model in nonmathematical terms as a way of illustrating his theoretical framework of *probabilistic functionalism*, which views the individual as a “natural statistician” who must adapt to a probabilistic world. A mathematical interpretation of Brunswik’s model was provided by Hursch et al. (1964), who modeled the individual and the environment as linear equations, allowed for error terms in both equations, and provided formulas that could be used to compute the decision maker’s quality (i.e., the match between predicted and real values). The lens model is customarily portrayed as in Figure 2, with the right-hand side representing the individual and the left-hand side representing the environment.\(^2\)

Brunswik’s model has influenced many empirical and theoretical studies on perception, learning, and decision making. Empirical work based on the lens model typically use experiments or field studies to identify factors that affect the accuracy of human judgment. For instance, Kessler and Ashton (1981) studied what type of feedback (i.e., relating to cues, weights, or outcomes)

\(^2\)It is customary to draw all of the arrows in the lens model’s graphical depiction, but this does not imply that the individual pays attention to every environmental cue. Cues that affect one side but not the other can have their $\beta$’s set to 0.
An important line of theoretical work has used the lens model to study when simple decision rules can perform nearly as well as more complex, optimally derived rules. Examples of the former rules are take-the-best (i.e., deciding based simply on the most relevant cue; Gigerenzer and Goldstein 1996) and equal weights (i.e., deciding based on an average of the cue values; Einhorn and Hogarth 1975). In terms of the lens model, take-the-best corresponds to setting the most relevant cue’s $\beta^\text{Ind}_1$ to 1 and the others to 0; and equal weights corresponds to setting all the $\beta^\text{Ind}_i$ to $1/N$. Hogarth and Karelaia (2007) study the effectiveness of different decision rules (including take-the-best and equal weights) in simulated environments and find that a given rule’s effectiveness is determined by its fit with the environment. For instance, in environments where a single cue is especially relevant, take-the-best will perform well; in contrast, if the contribution of cues is compensatory (i.e., if a low value on one cue can be compensated by a high value on another cue) then equal weights will perform well.

2.3 Combining Brunswik’s Lens Model and Information Aggregation

There are several good reasons to combine Brunswik’s lens model with information aggregation. First, doing so facilitates bridging the gulf between executive characteristics and organizational outcomes; this is because the lens model covers processes I and III, thus complementing process II. Specifically, under a model that integrates the lens model with aggregation, the elements
of Figure 1 correspond to the following: (a) Figure 1’s “executive characteristics” correspond to the cues and weights in the executives’ mental representations of the environment; (b) an individual’s cognition corresponds to the right-hand side of the lens model; (c) the aggregation structure is the decision rule that combines individuals’ opinions and produces a group-level decision; (d) the environment corresponds to the left-hand side of the lens model; and (e) the organizational outcome depends on the quality of the organization’s decisions (i.e., how appropriate are the organization’s decisions given the environment).

Second, the broad empirical support enjoyed by the lens model provides a solid foundation for models of aggregation. Because process II operates on the output of process I, using an empirically validated process I makes predictions about process II more reliable. The empirical nature of the lens model literature also suggests a clear path for future work in that it provides plenty of ideas about how to estimate cognition and the environment.

Finally, the lens model is a good candidate for providing a micro foundation for the behavioral theory of the firm. The lens model speaks to the cognitive aspects of that theory while sharing with it a similar sensibility: a focus on decision making, bounded rationality, and parsimonious process descriptions of behavior.3

In sum, melding the lens model and information aggregation is a valuable exercise because it allows one to bridge the gulf between executive characteristics and organizational outcomes and because it opens up a number of empirical and theoretical avenues. The aim of the ensuing model and analyses is to illustrate some of that value.

3It is interesting that Brunswik’s lens model shares common origins with the behavioral theory of the firm. With Administrative Science (Simon 1947/1997), Simon aimed to provide logical and psychological foundations to the field of organizations (p. xi). As his main influences toward these two goals he acknowledged Rudolph Carnap and Edward Tolman, respectively (Crowther-Heyck 2005:101). It turns out that Carnap was founder of the Vienna Circle, of which Brunswik was a member; and that Tolman was a close collaborator of Brunswik (Ash 2001).
3 Model

3.1 Overview

The model describes an organization whose behavior depends on the processes outlined in Figure 1. The organization in the model is tasked with screening a stream of projects—that is, approving the good projects and rejecting the rest. Screening projects is a common task in many settings, such as deciding whether to hire an employee, acquire a firm, or launch a new product.

A project is described by some characteristics or cues (in the context of a car company, for example, the cues could be the color and horsepower of a new car) and by a quality measure that depends in part on the value of the cues (e.g., the profits generated by a new car are to some extent a function of its color and horsepower). For clarity of exposition, the initial analyses assume that projects are adequately described by two cues (this assumption is relaxed in the Appendix).

The relationship between project cues and quality describes an environment. For example, an environment could be described in this way: \( \text{quality} = -5 + 3 \text{color} + 7 \text{horsepower} - 2 \text{color} \times \text{horsepower} + \varepsilon \) (using scales that are suitable for describing color and horsepower). Two main characteristics of the environment are studied, complexity and uncertainty. Complexity determines how strong are the interactions between project cues (i.e., it controls the coefficient for \( \text{color} \times \text{horsepower} \)); uncertainty determines how random is the environment (i.e., it controls the variability of \( \varepsilon \)).

The exact relationship characterizing the environment is unknown to individuals, who instead have only a mental representation of it. A mental representation captures the relationship between cues and quality as understood by an individual. Two types of individuals are studied, a specialist and a generalist. The specialist’s mental representation does not include all the cues in the task environment (e.g., a marketing manager may believe that quality depends only on the car’s color whereas an engineering manager may believe that quality depends only on its horsepower).\(^4\) In contrast, the generalist’s mental representation includes all of the terms in the

---

\(^4\)Division of labor is a process that can lead to a specialist’s mental representation. For instance, a brain surgeon may not be attuned to symptoms indicative of the common flu.
environment (i.e., not only the main effects but also the interaction term).

Individuals infer the coefficients for their mental representations from projects that they have previously observed. So based on those projects, a marketing manager may believe that profits vary as a function of color according to (say) $-6 + 4\text{color}$; similarly, an engineering manager may believe that profits vary as a function of horsepower according to $-3+5\text{horsepower}$. An individual’s experience corresponds to the number of past projects that she has observed. An individual’s opinion about a new project is based on her mental representation—as when (continuing the example) the engineering manager’s opinion is that a car with a horsepower of 1 will generate profits of $-3 + 5 = 2$.

The main aggregation structure studied in this paper averaging, under which the organization makes decisions based on the average opinion of two specialists (each one specialized on a different cue). This structure is representative of firms with co-CEOs or with a CEO who makes decisions based on the opinion of two subordinates. In this paper, the performance of averaging is contrasted with the performance of letting just one individual (either a generalist or a specialist) make decisions for the organization. Letting one individual dictate the organization’s decisions is the simplest possible aggregation structure, and it is representative of a firm managed by an autocratic CEO or of a start-up managed by a lone entrepreneur. Performance is measured in terms of the average quality of the projects approved by the organization.

The model presented here captures the three processes depicted in Figure 1 as follows. Cognition is modeled by individuals who have different types of mental representations and experience. Aggregation is modeled in terms of different decision-making rules (i.e., averaging the opinion of two managers versus letting one manager dictate the firm decisions). And the environment is modeled in terms of the projects—of varying complexity and uncertainty—that the organization contemplates undertaking. This model is used to study how organizational performance depends on the characteristics of these three processes. The rest of this section gives a formal description of the model’s elements.
3.2 The Environment

A project has multiple characteristics or cues, \( x_1, \ldots, x_N \), corresponding to project aspects that are visible to managers. To simplify matters, in most of the paper it is assumed that \( N = 2 \) (the case of more cues is analyzed in the Appendix). Continuing with the car example, if \( x_1 = 0 \) and \( x_1 = 1 \) represent (respectively) blue and red and if \( x_2 = 0 \) and \( x_2 = 1 \) likewise represent low and high horsepower, then the project described by \( x_1 = 0.9 \) and \( x_2 = 0.7 \) would be a reddish car with a fairly powerful engine.

Apart from the cues \( x_1 \) and \( x_2 \), a project also has a quality \( y \). Quality depends on the cues according to the following equation (which is unknown to the model’s actors):

\[
y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon. \tag{1}
\]

This equation describes the environment in which the organization operates. In the case of a car company, for instance, Equation (1) describes how profits (\( y \)) depend on car characteristics (\( x_1 \) and \( x_2 \)), on exogenously given characteristics of the environment (the \( \beta \)'s), and on some error due to unpredictable factors (\( \epsilon \)) such as technological uncertainties or unpredictable competitors. This equation corresponds to the left-hand side of the lens model (cf. Figure 2) in that it establishes the true relationship between cues and outcome. Equation (1) incorporates an interaction term because interactions are a fundamental aspect of many organizational settings (Siggelkow 2011).

This paper explores the effects of two characteristics of the environment: uncertainty and complexity. Uncertainty (\( U \)) determines the variability of the errors, so that \( \epsilon \sim N(0, U) \) (where \( \epsilon \) is randomly drawn per project). Thus, as \( U \) increases, the relationship between the project characteristics (\( x \)'s) and the project quality (\( y \)) becomes more random. Complexity (\( K \)) determines distribution from which \( \beta_3 \) is drawn, so that \( \beta_3 \sim N(0, K) \) (with \( \beta_3 \) randomly drawn per environment). Thus, as \( K \) increases it is more likely that the interaction term will have a greater effect (either positive or negative) on the project’s quality.

The remaining variables in Equation (1) are not the focus of this analysis, and they are all randomly drawn from a \( N(0, 1) \) distribution (\( x_1 \) and \( x_2 \) are drawn per project; \( \beta_1 \) and \( \beta_2 \) are
drawn per environment). To study the effect of $K$ and $U$ (and not of a given random draw of any of the model parameters), the model is simulated many times, after which average performance can be reported conditional on the values of $K$ and $U$.\(^5\)

### 3.3 Individual Cognition

In the spirit of the literatures on bounded rationality and the lens model, the individuals in this paper’s model need not account for all project cues. One such individual is the specialist, whose mental representation accounts for just a single cue. The base model has two cues and so there are two specialists, $A$ and $B$, who specialize (respectively) on cues 1 and 2. Hence the two specialists respectively “think” the world behaves according to a mental representation of the following form:

$$y^A = \beta^A_0 + \beta^A_1 x_1 + \varepsilon^A \quad \text{and} \quad y^B = \beta^B_0 + \beta^B_2 x_2 + \varepsilon^B. \quad (2)$$

An individual uses the projects that she has seen in the past to estimate the coefficients of his mental representation. The number of previously seen projects—the individual’s experience—is denoted by $E$ (thus, a low $E$ corresponds to a novice manager and a high $E$ to a seasoned one).\(^6\)

Once these coefficients have been estimated, the mental representations of specialists can be written as

$$\hat{y}^A = \hat{\beta}^A_0 + \hat{\beta}^A_1 x_1 \quad \text{and} \quad \hat{y}^B = \hat{\beta}^B_0 + \hat{\beta}^B_2 x_2. \quad (3)$$

These equations correspond to the lens model’s right-hand side because each one establishes a relationship between cues and a predicted outcome. An individual can give her opinion about a proposed project $(x_1, x_2)$ by evaluating her mental representation—in other words, by plugging

\(^5\)The $x$’s in this model should be interpreted as normalized (scaled and centered) values representing the projects that enter the firm screening process. Hence, that organizations might screen only projects of a certain type (e.g., a car company that will consider making only high-horsepower cars) does not affect the range of the $x$’s because the normalization is applied to projects that actually entered the screening process. Restricting the type of projects screened may affect the $\beta$’s, though. For instance, if buyers of high-horsepower cars are particularly attracted to bright colors, then the interaction term $\beta_3$ may be large. The model allows studying cases such as this by varying $K$ and $U$.

\(^6\)Using past projects to estimate the coefficients of a mental representation is consistent with the interpreted signals framework (Hong and Page 2009:2180). Assuming a linear form for an individual’s mental representation is consistent with the Brunswikian approach described in Section 2.2 and also with the observation that linear models usually fit judgment data well (see Brehmer 1994 and the references therein).
the \((x_1, x_2)\) of the proposed project into Equation (3).

A generalist has a mental representation with the same form as the environment; that is,

\[
y^G = \hat{\beta}_0^G + \hat{\beta}_1^G x_1 + \hat{\beta}_2^G x_2 + \hat{\beta}_3^G x_1 x_2. \tag{4}
\]

The predictions of the generalist seldom correspond precisely to reality (i.e., normally \(\hat{y}^G \neq y\)), since it is unlikely that the \(\hat{\beta}^G\)’s estimated by the generalist will exactly match the \(\beta\)’s of the environment. In other words, the generalist’s mental representation is complete but not necessarily accurate.

### 3.4 Aggregation Structure and Organizational Performance

An aggregation structure transforms the opinions of one or more managers (i.e., \(\hat{y}\)’s) into an organization-level decision (e.g., a decision to approve or reject a project). Under the averaging structure, specialists \(A\) and \(B\) are consulted and the organization approves a project only if the average of their opinions is positive (i.e., approve only if \((\hat{y}^A + \hat{y}^B)/2 > 0\)). If an aggregation structure consults just one manager (i.e., one specialist or one generalist), then the resulting trivial structure is named as the corresponding manager. That is, the generalist “structure” approves a project if \(\hat{y}^G > 0\) and the specialist structure approves a project if \(\hat{y}^A > 0\) (or, equivalently, if \(\hat{y}^B > 0\)).

The performance of a structure is defined as the expected quality per screened project. That is, the sum of the qualities (the real \(y\)’s given by Equation 1) of the accepted projects divided by the number of screened projects. If quality is interpreted as profits, then performance amounts to the expected profits resulting from each structure.

The term \(\pi_s(K, U, E)\) denotes the performance of a given structure \(s\) (where \(s\) could be averaging, generalist, or specialist) in a given environment (determined by \(K\) and \(U\)) while employing individuals with experience \(E\). There is no closed-form expression for computing this measure,\(^7\) so performance is computed via simulation as follows. (i) Given an environment

---

\(^7\)The equations in this model may seem to be equivalent to those described in a linear models course (e.g., Searle 1971). However, the literature on linear models has developed mainly asymptotic results (i.e., as \(E \to \infty\)) and few that depend explicitly on \(E\). Even if asymptotic results were used (at the expense of not studying the
For each individual (two specialists and one generalist) use Equation (1) to simulate \( E \) random projects and their qualities. (ii) Use these simulated projects to estimate the mental representations of each individual via ordinary least squares.\(^8\) (iii) Create a new random project and use Equation (3) to compute the managers’ prediction about it. (iv) For each aggregation structure, record the quality of the accepted project (or 0 if no project was accepted). Finally, (v) repeat steps (i)–(iv) 10 million times; doing so yields a \( p \)-value below 0.01 for all reported results. The average of the recorded qualities is then an accurate measure of \( \pi_s(K, U, E) \).

Finally, in order to make results comparable across environments and to simplify the interpretation of the findings, performance is scaled to fall in the range from 0 to 1. Here, 1 represents the best possible performance (i.e., approving all projects with \( y > 0 \) and rejecting the rest) while 0 represents the natural low-performance benchmark of simply approving all projects (i.e., a “lazy” screener who approves everything and thus performs no screening at all).\(^9\)

## 4 Results

The goal of this section is to study how the performance of the different aggregation structures (averaging, generalist, and specialist) depends on the managers’ experience (\( E \)) and the environment’s characteristics (\( K \) and \( U \)). To convey the results in an intuitive yet precise way, the presentation is organized around a series of figures. These figures were created after exhaustively exploring the model’s behavior and then selecting a reduced set of scenarios representative of that behavior overall.

The figures that follow use a common set of parameter values. Complexity (\( K \)) and uncertainty (\( U \)) take values in the range from 0.125 to 8. Recall that \( K \) and \( U \) affect the standard effect of parameter \( E \), the interaction term in the environment equation (Equation 1) leads to a probability distribution that cannot be integrated analytically.\(^8\) The effect of individuals whose estimation method is less efficient than OLS is equivalent to decreasing experience \( E \).

\(^9\)For instance, if firms in a given environment only received five projects with qualities \{-20, -10, 0, 10, 20\}, the lazy screener would approve all the projects, getting a performance of 0 \((= [-20 + -10 + 0 + 10 + 20]/5)\); while the perfect screener would only approve the good projects, getting a performance of 6 \((= [10 + 20]/5)\). Thus, a firm achieving a performance of, say, 3 (halfway between the minimum and maximum performance in that environment) would be assigned a scaled performance of 0.5. This type of scaling is common in research that compares the performance of different decision rules in different environments (see, e.g., Payne et al. 1993:91). Mathematically, scaled performance is defined as \( \frac{\pi(\cdot) - \pi_{\text{min}}}{\pi_{\text{max}} - \pi_{\text{min}}} \); for \( \pi_{\text{max}} = \int f_{x_1} f_{x_2} y I[y > 0] \, dx_1 \, dx_2 \) and \( \pi_{\text{min}} = \int f_{x_1} f_{x_2} y \, dx_1 \, dx_2 \), where \( f_{\cdot} \) denotes the probability distribution of the corresponding cue.
deviation of (respectively) \( \beta_3 \) and \( \varepsilon \) and that the remaining variables in the environment equation have a standard deviation of 1. Thus 0.125 and 8 represent effects that are, on average, 8 times smaller and larger than the effect of other elements in the environment.

Experience \( E \) is varied in the range from 5 to 640. High experience is represented by \( E = 640 \) because, at that level, increasing experience does not improve performance in over 99% of the simulations (i.e., \( E = 640 \) describes asymptotic behavior). Low experience is represented by \( E = 5 \) because this is practically the minimum number of observations needed to estimate the generalist’s mental representation (which has four \( \beta \)'s, so at least four observations are needed to estimate them).\(^{10}\)

Figure 3 illustrates the effect of experience on performance in a given environment \((K = 0.125 \text{ and } U = 2)\). A straightforward observation from this figure is that performance increases with experience. The reason is that more experience allows individuals to develop more accurate mental representations and thus to make better predictions. This finding is in line with empirical research (Barr et al. 1992, Tripsas and Gavetti 2000, Gary and Wood 2011) and theoretical research (Gavetti and Levinthal 2000, Denrell et al. 2004, Gavetti et al. 2005) documenting that more accurate mental representations lead to fewer decision errors. In the current model, more experience is never detrimental because the environment remains stable (i.e., the coefficients in Equation 1 do not change within a given simulation). Yet under environmental change the effect of more experience could be neutral or even negative (see, e.g., Siggelkow and Rivkin 2005, Posen and Levinthal 2012).

An interesting characteristic of Figure 3 is that the best aggregation structure is a function of experience: the generalist performs worst with low \( E \) but performs best with high \( E \). In general, determining which structure will be the best as a function of experience \((E)\) and the environment \((K,U)\) is not straightforward to answer, as there are some nuanced interactions among these variables. The answer to this question is built up progressively as follows. Section 4.1 focuses on experienced managers \((E = 640)\), Section 4.2 focuses on inexperienced managers \((E = 5)\), and Section 4.3 compares structures whose members have different levels of experience. The

\(^{10}\) Setting \( E = 5 \) rather than \( E = 4 \) does not qualitatively affect the results, but it does yield easy-to-read round numbers on the figure axes that signify values of \( E \).
robustness of the findings is discussed in the Appendix, which shows that the results are robust with respect to: the number of cues in the environment, the degree of the interactions in the environment, and the weights assigned to each specialist by the averaging structure.

4.1 Experienced Managers

The panels in Figure 4 show the effect of uncertainty \((U\) on the \(x\)-axes) and complexity \((K\) in each panel) on the performance of each structure under experienced managers \((E = 640)\). Analyzing these panels helps one to understand some of the mechanisms driving the results and to establish a baseline against which later results can be compared.

A first observation from Figure 4 is that the performance of each structure decreases with uncertainty. The reason is that, as uncertainty increases, past observations are less predictive of the future (each past observation contains more noise and less information and so is less representative of the environment’s true structure); this handicaps the manager’s ability to learn about the environment.

A second observation from this figure is that the best-performing structure is the generalist followed by averaging and then by the specialist. This ranking reflects the progressive oblivious-
Figure 4: Performance of aggregation structures under experienced managers ($E = 640$).

Figure 4 uses the same parameters as Figure 5, but now the managers in each structure have low experience ($E = 5$). A striking difference between Figure 4 and Figure 5 is that, in the latter, the generalist is not always the best-performing structure: panels (a) and (b) in Figure 5 show crossovers among the performance curves. The implication is that relying on the simple
mental representations that characterize the specialist and averaging structures may yield better performance than relying on the more complete mental representation used by the generalist.

To understand the conditions under which simple or complex mental representations are preferable, it is useful to start by analyzing the generalist. Figure 5c shows that the generalist is the best performer when complexity is high. The reason is that only the generalist structure “sees” the environmental interaction, which is consequential when complexity is high.

If complexity is not especially high, as in panels (a) and (b) of the figure, then accounting for the interaction is not consequential and instead uncertainty becomes the decisive factor. In these cases, the generalist continues to be the best performer when uncertainty is low because she can still estimate an accurate mental representation (at the extreme, if \( U = 0 \) then just four past observations are enough to estimate perfectly the four \( \hat{\beta}'s \) of the generalist).

But as uncertainty increases in panels (a) and (b), the generalist’s performance eventually falls below the performance of the other structures. This happens because, when experience is low, a generalist counts with little data to estimate accurately the many parameters in her mental representation and so may “overfit” these parameters to the few observations that she saw. This phenomenon is problematic when uncertainty is high, since it is unlikely that just a few observations will reflect the environment’s true structure. Therefore, an inexperienced generalist facing an uncertain environment will probably make faulty inferences and thus affect performance negatively.

Under these conditions, averaging has two advantages over the generalist. First, averaging
includes two individuals, which collectively have twice the experience of the generalist. Second, each of the individuals whose opinions are averaged estimates only two parameters (rather than the four parameters that must be estimated by the generalist); hence each has at her disposal twice as much data per parameter to be estimated. This second advantage applies also to the specialist, which is why—if uncertainty is high enough—even a single specialist outperforms the generalist in panels (a) and (b) of Figure 5.

A powerful way to understand when the generalist structure is preferable to averaging, and vice versa, is in terms of a bias–variance trade-off (Geman et al. 1992; see also Hastie et al. 2009). The generalist has a model that matches the true structure of the environment, and hence can estimate the environment with little bias if given enough past observations. In contrast, the specialists whose opinions are averaged do not consider the interaction and can therefore estimate only a biased model. Yet because they are two individuals with access to different pools of experience and because they estimate fewer parameters than the generalist, averaging their opinions incorporates more data per parameter and thus can estimate a model that has less variance. In other words, the generalist achieves lower bias but higher variance whereas averaging achieves lower variance but higher bias. The combined effect of these two sources of errors tilts in favor of averaging when complexity is not high (which keeps the bias of averaging at a moderate level) and also when uncertainty is high and managers are inexperienced (which makes the extra data available to the averaging structure particularly valuable). The bias–variance tradeoff among the structures is further discussed in Appendix A.4.

One application of the previous logic is to qualify the situations in which following “simple rules” may be a beneficial firm strategy (Eisenhardt and Sull 2001, Bingham and Eisenhardt 2011). Eisenhardt and Sull (2001:107) exhort managers to use a “few straightforward, hard-and-fast” rules and assert that simple rules are a source of competitive advantage. Their recommendation contrasts with that of Weick (1979:261), who encourages managers to “complicate yourself!”—that is, to use more complex mental representations in order to improve firm performance. Which advice should managers follow? In terms of the results discussed so far, these contrasting recommendations can be reconciled by considering the environments in which simple or complex decision-making rules are preferred. Averaging and specialists are aggregation struc-
tures that operate based on simple rules that do not take into account the environment’s full complexity. As explained above, these structures are more appropriate in environments characterized by low complexity, high uncertainty, and inexperienced managers (see the right halves of panels a and b in Figure 5). Arguably, these characteristics match those of the start-ups studied by the research finding that simple rules are beneficial to the firm (Eisenhardt and Sull 2001, Bingham and Eisenhardt 2011). In other environments, however, one might suppose that relatively complex rules have the upper hand.

4.3 The Equivalent Generalist

So far the analyses have assumed that all individuals have the same level of experience. In the real world, however, one may reasonably suppose that it is easier to hire experienced specialists than experienced generalists. Managers who are early in their careers acquire mostly functional training, which limits their opportunities to become generalists; later in their careers, the few select managers able to broaden their perspective (because, e.g., they have become CEOs) may never become experienced generalists owing to time and cognitive constraints. These dynamics should lead to a relative scarcity of experienced generalists.

In the rest of this section, the analysis assumes that averaging counts with experienced specialists (i.e., \( E = 640 \)) and then asks what experience a generalist should have in order to perform as well as averaging. For ease of the presentation, such a generalist will be called the equivalent generalist with experience level denoted \( E^G \).

Figure 6 displays the experience of the equivalent generalist \( (E^G) \) for a broad range of environments \((K \text{ varies along the } x\text{-axis and } U \text{ varies along the } y\text{-axis})\). The main observation derivable from this figure is that, in environments with high uncertainty but low complexity (roughly speaking, the upper left portion of the figure), very experienced generalists \((E^G \geq 80)\) are needed in order to match the performance of averaging with experienced specialists. As explained previously, high uncertainty makes learning harder for the generalist (because she has more parameters to fit and relatively less available data), while the absence of a high complexity environment does not make her richer mental representation particularly valuable. The high experience required by the generalist in Figure 6’s upper left contrasts with the rest of the
Figure 6: Equivalent generalist—the level of experience $E^G$ at which the generalist performs as well as averaging experienced individuals ($E = 640$).

In the figure, where generalists with little experience ($E^G \leq 10$) are able to match the performance of averaging experienced specialists.

One conjecture that emerges from this analysis is that, in uncertain environments that are not complex (i.e., the upper left of Figure 6), one should expect to see more firms managed by a process akin to averaging rather than by a generalist, as generalists with such high experience are likely to be scarce. That would explain why having co-CEOs may be beneficial for some firms. This may be the case with Whole Foods and Chipotle, both of which are successfully managed by co-CEOs. These businesses arguably face high uncertainty (they are among the first in their respective industries to serve a new demographic of health-oriented and environmentally conscious customers) but low complexity (given that the business of supermarkets and restaurants is both well understood and modular).

Another conjecture concerns the environments in which generalists with little experience outperform averaging experienced specialists. These environments are characterized either by low uncertainty or high complexity (roughly the cases where $E^G \leq 10$ in Figure 6). In such environments, one should expect to see most firms being managed by generalists. In other words, among firms facing low uncertainty (e.g., utilities, monopolies) or high complexity (e.g., firms
developing new technological ecosystems), even inexperienced generalists should perform better than other aggregation structures.

Steve Jobs and the computer industry of the 1980s helps to illustrate this latter conjecture. A computer’s value depends in a complex way on both technological and aesthetic factors. Apple Computer may have flourished under Steve Jobs because he understood how technology and aesthetics interacted whereas most of his peers (who were chiefly technologists) did not. Because complex environments are where the inexperienced generalist outperforms the other structures, it should not be surprising that Apple’s performance deteriorated sharply after the CEO baton was passed from Jobs (then, an inexperienced generalist) to John Sculley (a marketing specialist).

5 Discussion

This section discusses the practical and theoretical implications of the results presented so far.

5.1 Implications for Practice

In an ideal world, firms would be better managed by experienced generalists—individuals whose mental representation is both complete and accurate and who can thus perfectly apprehend the environment. Yet in the real world, such individuals are scarce or nonexistent. One must therefore ask whether is it better for the firm to be managed by an inexperienced generalist or by a structure that uses experienced specialists—individuals who, despite having an incomplete mental representation, have a better understanding of those (fewer) environmental aspects to which they attend.

Inexperienced generalists are preferable over averaging the opinions of experienced specialists when the environment is characterized by either low uncertainty or high complexity (i.e., roughly the lower and right portions of Figure 6). When uncertainty is low, a generalist’s mental representation is well suited for accurately apprehending the environment based on relatively few observations. When complexity is high, the generalist’s unique ability to discern environmental interactions also gives her an advantage.

Arrington’s (2011) observation that the best Internet CEOs are inexperienced 25-year-olds
can be seen to follow from two factors. First, that particular industry is new and so there really are no experienced managers. Second, it is a complex industry and so those who are more likely to have developed a generalist mental representation are the ones approaching this industry with a *tabula rasa*; that is, they are not primed with a mental representation from some other industry featuring interactions that are not relevant to the Internet. Both factors are in line with Internet CEOs being young and inexperienced. But this need not be the case in the future: as the Internet ceases to be a “new” industry, some managers may acquire enough experience to become experienced generalists; or if complexity decreases, the industry may prove to be better managed by specialists.

Experienced specialists working in tandem are preferable to inexperienced generalists when the environment is uncertain but not too complex (as in the upper left of Figure 6). In such cases, the more complete mental representation of the generalist becomes a liability: she has more parameters to estimate but does not have access to much data and does not extract much value from her knowledge about the environment’s interactions. As mentioned in Section 4.3, this may explain why firms operating in uncertain but less complex environments—such as restaurant and supermarket chains serving new consumer segments—have succeeded while using co-CEOs.

The fact that in environments that are uncertain but not overly complex specialists can outperform generalists may give credibility to Lutz’s (2011) call for car companies to be managed by specialists (“car guys” as he calls them). Arguably, the car industry is not high complexity (as after over a century of development, product and service interfaces have become mostly modular) yet new developments such as driverless cars and ride-sharing services like Uber make the industry uncertain. Thus the car industry falls somewhere in the upper left corner of Figure 6, where specialists working in tandem can outperform very experienced generalists. Furthermore, if these new developments substantially devalue managers’ experience (i.e., making them effectively inexperienced), then car companies may fare better when managed by even a single specialist—as in the right halves of panels (a) and (b) in Figure 5—than by a generalist. Perhaps that is why the leading electric car company Tesla is managed by Elon Musk, a technology specialist.

One practical implication of special interest to academics concerns the type of professionals
that MBA programs should graduate. The MBA curriculum faces a breadth–depth trade-off. That is, in light of constraints on time and learning, an MBA program can graduate either students who know broadly about many different business aspects or students who understand relatively fewer aspects but in much more detail (in fact, MBA programs differentiate along these very dimensions: some programs focus on general management whereas others focus on functional specialties). In terms of this paper, MBA programs can either graduate inexperienced generalists or experienced specialists. On which of these two profiles should an MBA program focus? The current paper has shown that both profiles are valuable under different environments. Thus, an MBA program that seeks to increase student placement should attempt to forecast the environment its graduates will face and then focus the program on the profile that will be most in demand. So if uncertainty is expected to increase (i.e., leading to an environment near the upper left of Figure 6), then the MBA program should prepare more specialists than generalists; it should also impart teamwork skills, since specialists add the most value when they are working with other specialists.

5.2 Implications for Research

The goal of this research has been to develop a theory that would bridge the gulf between executive characteristics and organizational outcomes. This paper underscores that the elements of that bridge—cognition, aggregation, and the environment—are deeply intertwined. It is therefore crucial for research on organizations to study these three processes together, or at least to control for processes that are not the focus of a given study. Otherwise, research in this field will risk perpetrating an “aggregation fallacy.” The organizations field cannot simply assume that results from psychology (which speak to process I) will be directly reflected as organization-level outcomes; similarly, it cannot assume that results from economics (which speak to process III) will play out irrespective of the firm’s structural characteristics and of its managers’ cognitive characteristics. What is distinctly organizational—namely, the aggregation process—must be at the core of research on organizations.

A general implication of the research presented in this paper is that, in some situations (e.g., environments characterized by the upper left area in Figure 6), the aggregation structure
can be used to compensate for incomplete and inaccurate mental representations; in particular, process II may compensate for flaws in process I. Suppose, for instance, that all individuals are specialists and therefore sense only partial aspects of the environment. Within the right aggregation structure, these individuals might still enable the organization to behave as if it perceives the whole environment. Thus the aggregation structure may allow the proverbial blind men to “see” the elephant.

This implication is in line with the idea espoused of the Carnegie tradition that, although individuals possess limited cognitive capabilities, organizations can overcome those limits by designing appropriate decision-making structures (Simon 1947/1997:92–93). To study that idea rigorously, this paper has incorporated Brunswik’s lens model into a model of organizational decision making. Using Brunswik’s model in this way helps to answer the call by Gavetti et al. (2007:530) for expanding the behavioral theory of the firm by incorporating core developments from psychology.

For the sake of clarity, the framework presented in this paper is deliberately simple. This means that much empirical and theoretical work remains to be done in describing the characteristics and interactions among the processes that “bridge the gulf.” Further empirical work could use experiments to measure how cognition, aggregation, and the environment jointly affect organizational performance. For this it would be critical to build on the rich empirical literature based on Brunswik’s lens model (for references, see Cooksey 1996, Hammond and Stewart 2001, Karelaia and Hogarth 2008). One way that the organizations field can add value to this vast body of work is by embedding Brunswik’s lens model into realistic organizational settings—that is, by considering the effect of the aggregation structures and environments that executives must deal with. Further theoretical work could contribute richer models of both the environment and the aggregation structure. For instance, it could study environments in which some cues are consistently more relevant than others (and ask, e.g., when is it preferable not to include specialists on the less relevant cues) as well as aggregation structures that are more complex (in that, e.g., they include multiple decision steps or decision makers).

It is possible for organizations whose members have inadequate mental representations to make good decisions; conversely, organizations whose members have complete and accurate
mental representations can nonetheless make bad decisions. An extreme case, described by Weick (1995:54–55), is that of a group of exhausted soldiers who were lost in the Alps yet escaped from their ordeal despite having the wrong map. At the other extreme, Janis (1972) describes the case of President Kennedy’s cabinet during the Bay of Pigs invasion: most of them had correct information about the low chances of success, but still they did not call off the invasion. The research reported here creates a bridge between the characteristics of these decision makers and the performance of their organizations. Understanding the processes underlying this bridge may help organizations make the most of their circumstances. It may also help in the quest to uncover the micro-to-macro processes that underpin organizations.
References


A Appendix: Robustness Checks

This appendix checks the robustness of the reported findings with respect to three assumptions of the model: (i) the number of cues in the environment, (ii) the degree of the interactions in the environment, and (iii) the weights assigned to each specialist by the averaging structure. In addition, a bias–variance decomposition is used to further explore the robustness of the results.

A.1 Increasing the Number of Cues

For the sake of simplicity, the analyses so far have assumed that projects have $N = 2$ cues, $x_1$ and $x_2$. One way to justify this assumption is in terms of an environment with larger $N$ but in which two cues are clearly the most relevant (i.e., whose $\beta$’s have the greatest magnitudes). Alternatively, the two cues could be the most representative of other cues’ behavior—as when there is a correlation structure among the $N$ cues and $x_1, x_2$ are the principal components of that structure.

Although there are cases in which $N = 2$ is justifiable, there are doubtless many situations characterized by $N > 2$. In fact, 176 of the 249 lens model studies analyzed by Karelaia and Hogarth (2008) had three or more cues. This section explores whether the results presented in this paper continue to hold when $N > 2$.

Modifying the model to accommodate more cues implies changing the equations for describing the environment and mental representations. For instance, the environment equation for an $N = 4$ case with all two-way interactions present is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1 x_2 + \beta_6 x_1 x_3 + \beta_7 x_1 x_4 + \beta_8 x_2 x_3 + \beta_9 x_2 x_4 + \beta_{10} x_3 x_4 + \varepsilon. \quad (5)$$

Note that increasing $N$ from 2 to 4 increases the number of two-way interactions from 1 to 6 (in general, the number of two-way interactions is $\binom{N}{2} = N(N - 1)/2$).\(^\text{11}\)

When $N = 4$, the mental representations of the generalist and the specialists are defined as follows. The generalist’s mental representation has the same form as Equation (5), but the

\(^{11}\)An alternative model could add a complexity parameter $K$ to control for the percentage of two-way interactions that are nonzero. The effect of such a parameter would be similar to that of adding higher-order interactions (discussed next).
specialists’ can now take multiple forms. For instance, apart from the one-cue specialists defined in Equation (3), one could define two-cue specialists with mental representations

$$\hat{y}^A = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 \quad \text{and} \quad \hat{y}^B = \hat{\beta}_0 + \hat{\beta}_3 x_3 + \hat{\beta}_4 x_4.$$  

Figure A.1 shows the effect of experience on performance in an $N = 4$ environment; panel (a) assumes one-cue specialists and panel (b) assumes two-cue specialists. The figure is set in the same environment as Figure 3 (i.e., $K = 0.125$ and $U = 2$). The performance of the specialist and of averaging increases as one moves from panel (a) to panel (b) because the respective mental representations account for an increasingly larger share of the environment’s variance. Hence the effect of increasing the number of cues to which a specialist attends is similar to the effect of keeping the number of cues constant while decreasing uncertainty; thus ignorance and uncertainty have equivalent effects.
A.2 Including Higher-Order Interactions

Until this point, the only interactions in the environment have been two-way interactions (i.e., the $x_ix_j$ in Equations 1 and 5). Of course, there also exist environments with higher-order interactions. Figure A.2 analyzes one such environment; it has $N = 4$ cues, all possible two-way interactions, and all possible three-way interactions (i.e., Equation 5 + $\beta_{11}x_1x_2x_3 + \beta_{12}x_1x_2x_4 + \beta_{13}x_1x_3x_4 + \beta_{14}x_2x_3x_4$). In this figure, the specialists used by averaging are one-cue specialists (as in Figure A.1a).

One difference between Figure A.1a and Figure A.2 is that, in the former, the generalist’s performance ramps up faster: the generalist surpasses averaging at about $E = 20$ in Figure A.1a but does so at about $E = 30$ in Figure A.2. The reason for this difference is that the generalist in Figure A.1a has a simpler mental representation (one that does not include terms for the three-way interactions) and therefore requires less experience to surpass the performance of averaging.

In general, the effect of adding higher-order interactions to the environment depends on
experience. Low levels of experience increase the advantage of averaging as compared with the
generalist; this is because, in a more complex environment, the inexperienced generalist has a
more difficult learning task. Yet high levels of experience increase the generalist’s advantage over
averaging because, in a more complex environment, specialists are aware of a smaller proportion
of environmental terms.

A.3 Averaging under Randomly Assigned Weights

The averaging structure puts equal weight on each specialist. Such a perfect balance may be
unlikely in the real world, where decision structures often place more weight on the opinions
of some specialists rather than others (reflecting, e.g., power or personality differences among
individuals). To account for that possibility, the analysis now turns to an averaging structure
that weights each individual’s opinion randomly. This randomly weighted averaging structure
approves a project only if \( \omega \hat{y}^A + (1 - \omega) \hat{y}^B > 0 \), where \( \omega \sim U[0, 1] \) (with \( \omega \) randomly drawn per
project).

Figure A.3 compares randomly weighted averaging to averaging under the same conditions
used in Figures 4 and 5. Randomly weighted averaging underperforms averaging because assign-
ing unequal weights to the specialists is inconsistent with the symmetric nature of the specialists
(i.e., each specialist knows about a different cue and, on average, all cues have the same rele-
vance).

An interesting aspect of Figure A.3 is that the performance loss due to random weighting is
quite small; in the worst case (Figure A.3a), the maximum performance loss is less than 5%. The
robustness of averaging to the choice of weights is in line with the literature on the robustness
of “improper” decision rules (see, e.g., Dawes 1979).

A.4 Decomposing Errors into Bias and Variance

Figure A.4 illustrates the arguments (discussed in Section 4.2) regarding bias and variance. Because the logic underlying the bias–variance decomposition depends neither on the number
of cues nor on the degree of the interactions, this analysis provides further evidence regarding
the robustness of the results.
Figure A.3: Averaging versus randomly weighted averaging. The upper and lower rows of panels employ the same conditions as in Figure 4 and Figure 5, respectively.
Panel (a) in Figure A.4 compares an inexperienced generalist (with \( E = 5 \)) to averaging experienced specialists (with \( E = 640 \)) in an environment of intermediate uncertainty \( (U = 1) \) while varying complexity \( (K \) on the \( x \)-axis). The \( y \)-axis in this panel is the mean squared error (MSE) of the opinions of the generalist and averaging (i.e., respectively \( \mathbb{E}[(\hat{y}^G - y)^2] \) and \( \mathbb{E}[(\hat{y}^A + \hat{y}^B)/2 - y)^2] \)). Panel (a) reveals that, when complexity is high (roughly, \( K > 3 \)), averaging makes a larger estimation error than the generalist. This result is consistent with the right side of Figure 6, which shows that \( E = 5 \) is enough experience for a generalist to match the performance of averaging experienced specialists.

Panels (b) and (c) in the figure show how the two sources of errors, bias and variance, affect (respectively) the generalist and averaging. These panels decompose the MSE of the two structures shown in panel (a) into its bias and variance components (Geman et al. 1992). In accordance with Section 4.2’s arguments on bias and variance, the two structures achieve their respective MSEs in opposite ways: panel (b) shows that the MSE of the inexperienced generalist is due mostly to variance, and panel (c) shows that the MSE of averaging experienced specialists is due mostly to bias.