A Structured Approach to Strategic Decisions
Reducing errors in judgment requires a disciplined process.
BY DANIEL KAHNEMAN, DAN LOVALLO, AND OLIVIER SIBONY

Envision the following situations: A board of directors considers acquiring a competitor. A marketing team decides whether to launch a new product. A venture capital investment committee chooses among an array of startups to fund.

All those strategic decisions share a common feature: They are *evaluative judgments*. To make such tough calls, people must boil down a large amount of complex information to either

(1) numerical scores for competing options or (2) a yes-no decision on whether to choose a specific path. Of course, some management decisions are made without weighing quite so much information. But strategic decisions tend to involve the distillation of complexity into a single path forward.

Given how unreliable human judgment is, all evaluations are susceptible to errors. These errors can stem from known cognitive biases — or they can be random errors, sometimes called "noise." Unreliability in judgment has long been recognized and studied, particularly in the context of decision-making about hiring. We draw inspiration from that body of research and experience to suggest a practical, broadly applicable approach to reducing errors in strategic decision-making. We call it the Mediating Assessments Protocol (MAP), and we'll describe it here, after discussing the underpinning research.

Strategic Options Are Like Job Candidates
The body of research on the job interview, the most common tool for employee selection, contains a wealth of information about the accuracy of evaluative judgments.[1] Most companies still use traditional (unstructured) interviews to make a global evaluation. The interviewer starts with an open mind, accumulates information about the candidate, and then reaches a conclusion.

Unfortunately, a vast amount of evidence indicates that unstructured interviews lead to biased evaluations that have very little predictive value.[2] That's because the interviewer forms a *mental model* (colloquially known as an "impression") of a candidate, a process that psychologists have shown has three specific limitations:

1.   **Excessive coherence.** Mental models are usually simpler and more coherent than the reality they aim to assess. As interviewers, if we assume, for instance, that a particular candidate is an extrovert, we tend to ask questions that confirm this hypothesis.

2.   **A "quick and sticky" quality.** We form our mental models rapidly, often on the basis of limited evidence at the start of the process, and we alter our models slowly as new facts emerge. That explains why, as common sense would suggest (and research has confirmed[3]), first impressions have a disproportionate effect on the assessments we make of people in general and on the outcome of job interviews.[4]

3.   **Biased weighting.** Our mental models often don't give each pertinent fact the weight it deserves. We may discount important bits of information or, in contrast, give great weight to factors that should be entirely

irrelevant. For example, an interviewer may wrongly perceive that a male candidate has great leadership qualities just because he is tall and has a deep voice.

Such challenges in hiring are easy to recognize. For that reason, we do not expect all interviewers to agree on one candidate — and we often compensate by averaging several interviewers' viewpoints.

What many people do not fully appreciate is that we use the same process of mental model formation in strategic decision-making, with similar

limitations and results.[5]

Suppose, for instance, that you're selecting a location for a new plant. You must assess multiple factors: labor costs, technical feasibility, regulatory requirements, political stability in various regions, and so on. You already have a mental image of the candidate countries and cities. As you learn new facts about each prospective site, the bias toward excessive coherence leads you to confirm that image, which is likely to be much less nuanced and less ambiguous than the reality.

This bias may be even stronger if a team is collaborating on the decision:[6] As a favorite plant site emerges, its perceived benefits are exaggerated and its perceived costs underestimated. Then, when the team reviews the merits of possible locations, early impressions formed in the first minutes of the discussion are likely to weigh heavily on the final decision. Once an impression is formed, you will tend to ask leading questions that support your early views. Because of confirmation bias,[7] you will interpret ambiguous facts in light of preexisting attitudes. It takes a considerable amount of evidence to reverse an erroneous initial judgment.

Finally, just as a candidate's physical appearance may influence a hiring decision, certain attributes of a manufacturing site may carry undue weight. Predictably, extra weight is often given to recent and salient information (a case of "availability bias").[8] For example, you may overreact to recent news of political turmoil or overemphasize shortterm considerations. And given that most of us tend to be optimistic and overconfident in our forecasts, you may underestimate technical challenges in constructing the new plant.

In short, unstructured decision-making, whether in job interviews or in other, more strategic decisions, is vulnerable to both bias and "noise." The presence of noise explains why researchers find large variations whenever they systematically investigate the reliability of judgment.[9]

What We Can Learn From Structured Interviewing

Dozens of studies on personnel selection have shown conclusively that decisions are more accurate when interviews are structured rather than unstructured.[10] Therefore, a growing number of organizations, especially those that put a high premium on the quality of the talent they hire, have adopted structured interviews.

An early form of structured interviewing was developed in 1956 by Daniel Kahneman while he served in the Israeli Army, where he observed that holistic ratings given by interviewers were poor predictors of the future success of recruits. He replaced these ratings with separate scores on six attributes: sense of duty, sociability, energy level, punctuality, capacity for independent thought, and what was then called "masculine pride." A simple average of these scores proved to predict overall performance more accurately than did an intuitive evaluation based on an unstructured interview. An intuitive prediction made *after* these separate structured-

interview ratings were assigned was also useful; the combination of the two was the best performance predictor of all.

The significance of this finding, which Kahneman reiterated many years later,[11] is that the accuracy of intuitive judgment is much improved by delaying a global evaluation until the end of a structured process. The Israeli Army still uses structured interviewing today, as do companies such as Google, Amazon, and McKinsey.[12]

In a structured interview, the interviewer must rate several key traits before making a final evaluation. Scores on each attribute serve as *mediating assessments*: intermediary ratings, produced in a predetermined, standardized manner in order to be as fact-based as possible. The final evaluations are then derived from these ratings.

Surprisingly, although structured interviews are becoming increasingly common in hiring, the type of structured decision-making they exemplify is rare in other contexts. Strategic decisions typically involve a holistic evaluation of all available information, without explicitly scoring key attributes. We think structure and mediating assessments can improve the quality of those decisions, as well.

Core Elements of
Structured Decisions

MAP is a structured approach to grounding strategic decisions, like structured interviews, on mediating assessments. MAP has three core elements:

• **Define the assessments in advance.** The decision maker must identify a handful of mediating assessments, key attributes that are critical to the evaluation. In the decision to acquire a company, for example, the assessments could include anticipated revenue synergies or qualifications of the management team. This process is similar to one a hiring committee would follow when creating a job description that outlines attributes required for success in the position.

• **Use fact-based, independently made assessments.** People who weigh in on one aspect of a strategic option should not be influenced by one another — or by other dimensions of the option. Their opinions should be grounded in the evidence available. This approach is comparable to a wellorganized structured-interview process, in which job seekers are scored on each key attribute solely on the basis of their answers to relevant questions, calibrated using predefined scales.

• **Make the final evaluation when the mediating assessments are complete.** Unless a deal- breaker fact is uncovered (for instance, evidence of accounting fraud at the acquisition target), the final decision should be discussed only when all key attributes have been scored and a complete profile of assessments is available. This is similar to having a hiring committee review all the evaluations made by each interviewer on each key requirement of the job description before making a decision on a candidate.

The use of mediating assessments reduces variability in decision-making because it seeks to address the limitations of mental model formation, even though it cannot eliminate them entirely. By delineating the assessments clearly and in a factbased, independent manner, and delaying final judgment until all assessments are finished, MAP tempers the effects of bias and increases the transparency of the process, as all the assessments are presented at one time to all decision makers. For example, because salient or recent pieces of information are not given undue weight, the process preempts the availability bias. MAP also

reduces the risk that a solution will be judged by its similarity with known categories or stereotypes (an error arising from the representativeness bias[13]). When differentiated, independent facts are clearly laid out, logical errors are less likely.

Some decision makers will have an initial dislike for MAP, just as many recruiters still resist structured interviewing. The requirements may appear mechanical, and the limits it places on the role of intuition will not appeal to leaders who have been rewarded for "trusting their gut." Structured decision- making, based on mediating assessments, will be adopted only if it is viewed as offering a substantial improvement in decision-making quality.

Accordingly, we next examine MAP's application and benefits in two types of strategy decisions: large one-off decisions made by teams of executives or directors, and recurrent decisions made as part of formalized processes that, in aggregate, shape a company's strategy.

Structuring One-Off Strategic Decisions

Let's return to the example of a large company that is contemplating a major acquisition. A deal team, assembled to perform due diligence, begins by identifying key issues — revenue and cost synergies, hidden liabilities, cultural compatibility, and so on. It prepares a report to the board, with chapters that cover each issue. In the subsequent board meeting, the chair leads a thorough discussion in which members state their view of the merger and explain their reasoning. Questions are asked and answered, and the concerns that emerge are discussed in depth. Eventually, a vote is taken.

This approach is thorough and professional and seems unobjectionable. However, it resembles unstructured interviewing in a critical respect: The conclusion is drawn directly from a global image of the case, an image that emerges spontaneously and gradually as information is considered. In contrast, structured decision-making requires leaders to make separate, explicit assessments of each aspect and to use those assessments as the basis for a decision. In the case of the possible acquisition, it would proceed as follows:

•   Due diligence is completed as usual. If the deal team has done a good job, the key issues it has identified correspond to the required mediating assessments. The only novelty is that the deal team is required to conclude each chapter of its report with a summary rating. Each rating answers the question: "Leaving aside how important this topic is to the overall decision, how strongly does the evidence in this chapter argue for or against the deal?" Like the scoring of specific traits in a structured interview, each of these ratings should be based solely on the facts contained in the relevant chapter, not on data gathered about other topics.

•   The agenda of the board meeting follows the chapters of the report. Topics are reviewed in sequence (unless a deal-breaker fact emerges that ends the discussion). Drawing on the data supplied in each chapter, the board members, first individually and then as a group, consider whether to accept or amend the deal team's proposed rating on each assessment. In a hiring decision, the corresponding step would be for the hiring committee to review interviewers' scores on each requirement of the job description.

•   Board members are asked to refrain from making general comments about the acquisition until the mediating assessments are complete. At that point, ratings on all assessments are displayed. Only then does the discussion turn to the holistic evaluation of the deal. Eventually, a vote is taken.

As this example illustrates, the implementation of MAP does not substantially increase the overall burden of deciding. Trivial extra effort is required from the deal team to generate a summary assessment on each due diligence topic. A board discussion that is structured around mediating assessments will be more organized and focused than the usual process, but not necessarily longer or more contentious.

At this moderate cost, the rigor of formal structure in strategic decision-making has the benefit of sequencing the process such that important facts are less likely to be overlooked and thoughtful, selfcritical consideration of trade-offs is much more likely to occur. In this way, using MAP triggers conscious reflection. In typical unstructured decision- making, by contrast, we (often unconsciously) weigh losses more than gains, the near future more than the distant future, and vividly presented anecdotes more
than dull statistics.[14]

For instance, in the sequence of presentations and review meetings that lead up to an acquisition, the topics discussed typically change as new information becomes available. As urgency mounts and enthusiasm about the one-off deal builds, serious issues may be ignored or relegated, say, to an appendix of a presentation document. Meanwhile, the target's most appealing features gain increasing prominence, and the narrative rationale for the deal gets more and more compelling. It is easy to see how these dynamics can lead to mistakes in judgment. Using a structured decision-making process ensures that the checklist of key questions is defined in advance and that, at the time of final decision, all its elements are given sufficient visibility. This rigor limits the risk that a compelling narrative will sway the board.

Our advice is not to assign explicit weights to assessments, as some decision theorists would suggest. Executives reject that type of framework because it mechanizes decisions and removes room for intuition, including the possible interactions among the criteria. Unlike algorithmic decisionmaking, which aims to take subjectivity out of the decision entirely, MAP values intuition, provided that it is *informed*. The holistic judgment of experienced executives is valuable, but it must first be prepared by a profile of mediating assessments.

Nor do we advise treating all assessments equally. Rather, some dimensions that are simply more important than others should be framed as make-or-break assessments and evaluated first. For example, when evaluating a new technology, considerations such as time to market and cost are important, but the assessment of whether the technology will work at all is paramount.

Structuring Recurrent Decisions

Recurrent decisions, in aggregate, also produce strategic outcomes. Consider, for instance, launching new products in the fast-moving consumer-goods market, advancing products along an R&D pipeline in pharmaceuticals, or making small acquisitions in a broader roll-up strategy. Large organizations make countless decisions like these, and their collective impact on the business can be critical.[15]

MAP provides the same rigor and discipline to recurrent decisions as to one-off choices. In addition, when decisions of a particular type are repeated, a structured approach facilitates learning. Explicit mediating assessments make it easy to conduct a postmortem of past decisions, examine the reasoning and use of facts, learn which assessments may have been misguided, and revise the assessments or the scales accordingly. The

quality of judgments is likely to improve further if the assessments and the evaluation are expressed on a relative scale. (See "Using Percentiles as a Rating Scale," p. 72.)

MAP's most important contribution to routine decisions is in providing true standardization. People who have the same role in an organization and make similar decisions are assumed to be interchangeable, but companies do not normally check whether this assumption is correct. However, when organizations test it through a "noise audit," evaluations often differ by 50% or more between supposedly interchangeable professionals (including highly experienced ones).[16] This level of noise, which is ordinarily invisible, is obviously unacceptable.

To better understand how to structure recurrent decisions, consider the case of a venture capital firm we will call VC Co., which has redesigned its investment process. Like many investors, VC Co. followed a thorough evaluation protocol, focusing its attention in every case on the specifics that appeared most relevant to form a mental model of the investment target. To improve the quality of its decisions, the firm decided to adopt a structured decision approach. In doing so, VC Co. incorporated the key features of MAP.

First, to predefine the critical assessments, VC Co. engaged in a thorough discussion of the following question: "What are the key factors we should consider when making investment decisions?" The output was a list of criteria that would not necessarily be identical for another venture capitalist, because they reflected the investment philosophy of this firm. For example, one of these criteria is a judgment on the target company's founding team. This judgment, in turn, breaks down into specific subassessments such as problem-solving skills, technical expertise, resilience, and open-mindedness. Such qualities can be assessed on the basis of interactions with the team, its track record, and reference checks.

Second, VC Co. now ensures that its assessments are made separately, do not influence one another, and are summarized in a percentile-score rating. Organizing the data collection to make assessments independent of one another (to the extent possible) minimizes the halo effect. For instance, having seen an impressive product demo, VC Co.'s investment professionals may be tempted to rate the management team's skills favorably. MAP guides them to treat these as separate judgments.

Assigning a numerical rating to each assessment safeguards against the tendency to form overly coherent mental models. For instance, when VC Co. formulated its assessment of a target company's top team qualitatively, the wording was sometimes ambiguous: Depending on the mental model that had emerged from prior assessments, a word such as "strong" could be interpreted as either adding to a positive impression or expressing doubt. With a numerical rating, this risk is reduced.

Using percentile scores to make qualitative assessments puts recurrent decisions into context. To evaluate the caliber of a target company's founders, for instance, VC Co. now uses a "slate" showing the names of the dozens of founders to whom it has had exposure. Instead of asking whether a founder is "an A, a B, or a C," investment committee members ask: "On this particular subassessment — say, technical skills — is this person in the same league as Anna and the other five founders we regard as the very best on this dimension? Or is she comparable to Bob and the other entrepreneurs we view as belonging to the second quartile?" VC Co. applies the same comparative approach to other qualitative assessments, such as the potential defensibility of a company's competitive advantage.

Third, the company has formalized conduct of its decision meetings. Just as in one-off decisions, the meeting agenda follows the evaluation structure: VC Co. reviews each assessment separately before coming to a final decision. Summary narratives that encapsulate a rationale for the proposed investment are banned until all assessments have been reviewed one at a time and a profile of the ratings is displayed. A major benefit of applying MAP to a recurrent decision is that investment committee members have multiple chances to practice this new discipline and can continuously improve their skills and tools. For instance, each new evaluation of a founder adds a name to the slate, making comparisons with future candidates richer.

VC Co.'s early experience using MAP did not slow down the firm's decision-making. On the contrary, members of the firm have told us that the orderly flow of work in the new protocol actually saves time, and they have pointed to several investments where they believe it has led them to a distinctly better decision. Quantitative effects will take years to assess.

Using Percentiles as a Rating Scale

How should decision makers express their assessments? Standard rating scales with verbal anchors ("very good," "good," and so on) have the advantage of simplicity. However, the ambiguity of verbal labels is a major source of noise, because different people use different words to convey the same underlying judgment.

Percentile scales offer a possible solution. Many leading companies already use them to express judgments about the performance of employees on multiple dimensions. For example: "Jenny is in the top 10% of the population of junior executives for

raw intellect, but in the third quartile for interpersonal skills." We recommend extending the use of percentile scales to assessments in other domains. When evaluating the quality of a target company's go-to-market capabilities, for instance, a venture capital firm can ask: "How does this company compare with other companies in the same sector? Is it in the top 10%? In the top 25%?" The percentile scale requires a specified frame of reference, understood and shared by all. In this example, "all the potential targets we have evaluated" would be the most appropriate reference class.

Percentile scales have several advantages over other types of rating scales. First, they require the evaluator to bring comparable cases to mind and to think of the case at hand as one particular instance of a broader category. This approach, which has been called the outside view, is a powerful debiasing technique by itself.i

Second, percentile scales allow individual biases to be detected and corrected. For example, an overly lenient or optimistic person who rates 40% of cases as belonging to the top 10% will eventually be identified and trained to use the scale appropriately. When people

have learned to use a percentile scale well, they are said to be calibrated. Improving calibration is another major step in reducing noise in judgments.

Third, percentile ratings can easily be translated into policy. If underwriters, for instance, rate risks in percentiles, premiums can be priced on the basis of their ratings, and the company can decide at what percentile in the distribution of risks it sets the limit of what it is willing to underwrite.

Percentile scales can be challenging to define, introduce, and administer. But the accuracy gains they bring are worth the effort.

Whatever else it produces, any organization is a decision factory. Some of its decisions are made by people following clear rules. But many of the decisions that shape the future of organizations require time-consuming deliberation, analysis, and the balancing of multiple considerations. Such decisions cannot easily be "quality checked." To improve them, we must work on the processes by which they are made.

MAP is one way of doing so. By adding discipline to decision-making and limiting some well-known flaws, it brings quality assurance to complex decisions. While other decision support approaches, such as decision theory or advanced analytical models, share the same objective, MAP has some advantages. It is easy to learn, involves a minimal amount of additional work, and leaves senior executives some freedom to exercise intuitive judgment, albeit after a useful delay. As such, it should be a valuable tool for any leader who aims to raise the quality of an organization's decisions.

REFERENCES

1.   J. Levashina, C.J. Hartwell, F.P. Morgeson, and M.A. Campion, "The Structured Employment Interview: Narrative and Quantitative Review of the Research Literature," Personnel Psychology 67, no. 1 (February 2014): 241-293.

2.   J. Dana, R. Dawes, and N. Peterson, "Belief in the
Unstructured Interview: The Persistence of an Illusion," Judgment and Decision-Making 8, no. 5 (September 2013): 512-520; and D.A. Moore, "How to Improve the Accuracy and Reduce the Cost of Personnel Selection," California Management Review 60, no. 1 (November 2017): 8-17.

3.   M.R. Barrick, S.L. Dustin, T.L. Giluk, G.L. Stewart, et al., "Candidate Characteristics Driving Initial Impressions
During Rapport Building: Implications for Employment Interview Validity," Journal of Occupational and Organizational Psychology 85, no. 2 (June 2012): 330-352.

4.   C.Y. Olivola and A. Todorov, "Fooled by First Impressions? Reexamining the Diagnostic Value of Appearance- Based Inferences," Journal of Experimental Social Psychology 46, no. 2 (March 2010): 315-324.

5.   D. Kahneman, Thinking, Fast and Slow (New York: Farrar, Straus and Giroux, 2011).

6.   C.R. Sunstein and R. Hastie, Wiser: Getting Beyond Groupthink to Make Groups Smarter (Boston: Harvard Business Review Press, 2015).

7.   R.S. Nickerson, "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," Review of General Psychology 2, no. 2 (June 1998): 175-220.

8.   A. Tversky and D. Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," Science 185, no. 4157 (Sept. 27, 1974): 1,124-1,131.

9.   D. Kahneman, A.M. Rosenfield, L. Gandhi, and T. Blaser, "Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision-Making," Harvard Business Review 94, no. 10 (October 2016): 36-43.

10.   Levashina et al., "The Structured Employment Interview."

11.   Kahneman, Thinking, Fast and Slow.

12.   L. Bock, Work Rules! Insights From Inside Google That Will Transform How You Live and Lead (London: Hachette U.K., 2015).

13.   Tversky and Kahneman, "Judgment Under Uncertainty."

14.   Kahneman, Thinking, Fast and Slow.

15.   O. Sibony, D. Lovallo, and T.C. Powell, "Behavioral Strategy and the Strategic Decision Architecture of the Firm," California Management Review 59, no. 3 (May 2017): 5-21.

16.   Kahneman et al., "Noise."

i. D. Lovallo and D. Kahneman, "Delusions of Success: How Optimism Undermines Executives' Decisions," Harvard Business Review 81, no. 7 (July 2003): 56-63.