- Analysts must take particular care not to fall into the traps of satisficing, oversimplification, using out-of-date or mismatched information, and vagueness.
- Analysts should not expect a client to join them on the second floor of their argument if they have not built a firm conceptual foundation on the first floor.

CONSIDERING THE CASE STUDY

Review Case Study IV, "Financial Crises in the United States: Chronic or Avoidable?" Assume this case study was written in 2015 for generalists seeking to learn more about the US financial crisis that erupted in 2008.

- What was the primary thesis or argument of the paper?
- How does the author establish credibility in making this argument?
- Which schools does the author suggest provide the most useful insights in explaining what caused the financial crisis of 2008? Which of the five techniques does the author rely on the most to make the argument?
- To which of the seven common traps discussed in the chapter is the author most susceptible?

NOTES

1. William Bernbach, in *Hey, Whipple, Squeeze This*, ed. Luke Sullivan (Hoboken, NJ: John Wiley & Sons, 2008), 6.
2. William J. McGuire, "Personality and Attitude Change: An Information-Processing Theory," in *Psychological Foundations of Attitudes*, ed. A. G. Greenwald, T. C. Brock, and T. M. Ostrom (New York: Academic Press, 1968), 179–180.
3. Charles Mingus, "Creativity," *Mainliner 21* no. 5 (1977): 25, quoted by W. H. Starbuck and P. C. Nystrom, "Designing and Understanding Organizations," in *Handbook of Organizational Design*, vol.1 (New York: Oxford University Press, 1981), 9.

17

How Should I Portray Probability, Levels of Confidence, and Quantitative Data?

SETTING THE STAGE

Analysts owe each client their best judgments. After reviewing all the available data, reexamining the line of reasoning, and considering alternatives, analysts need to give their readers some idea of the likelihood that their analytic judgments will turn out to be correct. This is most often done with terms such as *likely, most likely,* and *almost certainly.* Analysts sometimes prefer to give a numerical percentage, such as *a 60 percent chance,* based on the strength of their data and the soundness of their reasoning.

In the wake of the 9/11 terrorist attacks, analysts have begun to document the level of confidence they have in their sources and in the accuracy of their judgments. This is a different process, and it often gets confused with the concept of probability.[1] Sometimes analysts confuse the difference between providing an assessment of a 60 percent chance of an event occurring and recording their levels of confidence in that judgment, which is based on the credibility of the available data and persuasiveness of the line of reasoning. For example, an analyst could be 70 percent confident in assessing that an event has a 60 percent chance of occurring based on incomplete data, or 95 percent confident in that same judgment that an event has a 60 percent chance of occurring based on a more robust set of evidence.

When presenting quantitative data and statistical analyses, analysts need to be particularly careful to avoid displaying the data in biased ways and to fully document how the figures were derived. Analysts need to apply their critical thinking skills when collecting data and assessing how accurately others present their data.

LOOKING MORE DEEPLY

One of the analyst's primary tasks is to assess human behavior. Such assessments reflect one's views of the strategic situation as well as the intellectual and emotional make-up of leaders and their supporters and adversaries. Analytic judgments take into account key players' motivations, ambitions, psychological strengths and weaknesses, and views on strategic issues—none of which can be measured precisely. Analysts must also consider organizational constraints and other influences on a leader's decision making to formulate their subjective judgments of the likely behavior of leaders or the direction of the national economy.

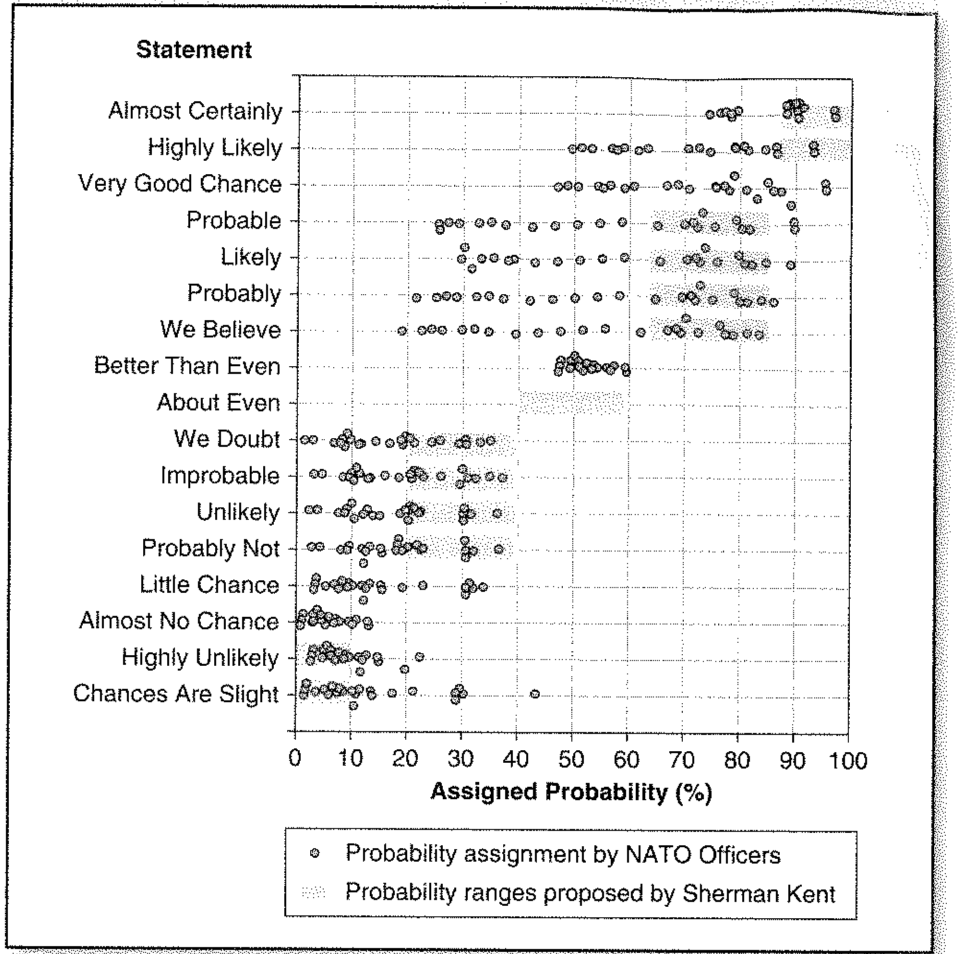
Because of all the uncertainties that must be factored into a product, analysts need to give their clients an overall assessment of the probability that they are correct. In providing these estimates, analysts can present either a verbal or numerical estimate. Verbal statements—such as *almost certainly, highly likely, probable, better than even, 50/50 chance*—give the readers some idea of the analysts' sense of the potential predictive accuracy of their judgments. Once again, analysts must avoid conflating the assessment of *it is highly probable that this will happen* with the concept of *it is highly probable that I am right that it is highly probable this will happen.*

Past experience—buttressed by research studies—cautions, however, that the meaning clients assign to such words of probability can vary substantially from what was intended. How clients interpret such words will be influenced by their own—and often differing—experiences with the use of such words as well as their personal preferences. In a study on measuring perceptions of uncertainty often cited in the Intelligence Community and summarized in Figure 17.1, NATO officers assigned a probability percentage to the word *probable* ranging from 25 to 90 percent. The same group of officers assessed the phrase *highly likely* as carrying a probability ranging from 50 to 95 percent. On the other end of the spectrum, the phrase *little chance* received probabilities ranging from 2 to 35 percent.

We have replicated this experiment hundreds of times in the classroom with the same results.[2] Invariably, students will allocate percentages ranging from 30 to 85 percent to the phrases *probably* and *most likely.* The wide range of percentages given would indicate that these words mean very different things to different people—raising the question of what information of value is actually being communicated.

Efforts to standardize the range of uncertainty associated with such words have met with limited success. Over a half-century ago, Sherman Kent proposed a schema for standardizing ranges of uncertainty that has yet to catch on as a universal standard despite its underlying logic (Kent's proposals are represented by the shaded boxes in Figure 17.1). Several US Intelligence Community agencies have tried to deal with this phenomenon by publishing tables of probability ranges similar to that generated by Kent, but these often are disregarded by the reader.

FIGURE 17.1 Measuring Perceptions of Uncertainty: NATO Officers and Sherman Kent



Source: Sherman Kent, "Words of Estimated Probability," in *Sherman Kent and the Board of National Estimates: Collected Essays*, ed. Donald P. Steury (CIA, Center for the Study of Intelligence, 1994); and *Scott Barclay et al, Handbook for Decision Analysis.* (McLean, VA: Decisions and Designs, Inc. 1977).

One of the difficulties can be demonstrated in how one interprets the word *probable.* In the classroom exercises cited previously, students were asked if the word *probable* can be used to represent a probability of less than 50 percent. In all instances, more than half of the class confidently answers "yes." Almost in every class, however, a small group of students—10 or 20 percent—answer just as vigorously "no." They explain that not all situations involve a choice between only two options. If, for example, five independent options are being considered and the likelihood of each option is 5, 20, 15, 35, and 25 percent, respectively, then the option with 35 percent likelihood should be considered the single most probable outcome. In this case, however, the outcome of "not the 35 percent outcome" is more probable than the 35 percent outcome or any of the other specific outcomes under consideration. We find that students often have difficulty differentiating between single most probable outcomes and most probable outcomes.

A further complication is that even if the producers of the analytic product reach consensus on what probabilistic language to employ, recipients of the document or briefing may either subconsciously or consciously interpret the phrase in a way more consistent with or supportive of their desired preference or outcome. For example, the client will often translate a term such as *likely* as meaning 70 or 80 percent if that is a desired outcome, but as only 50 or 60 percent if the outcome is not desired.

In the UK, analysts have been provided with an uncertainty yardstick to use in their intelligence products.[3] The yardstick arrays a probability range against a standardized qualitative term (see Figure 17.2). What is intriguing about the UK probability ranges is that a conscious decision was made to insert gaps between each level. The intent is to discourage analysts from trying to draw too fine a line in generating percentages and to remind recipients of their intelligence products that the terms are not intended to make precise distinctions.

Following the 9/11 attacks on the United States and the Iraq WMD fiasco, the National Intelligence Council (NIC) started publishing a probability chart showing the relative probabilities of key terms as depicted in Figure 17.3.[4] The NIC chart arrayed key terms from least to most likely without assigning specific percentages.

In 2010, the Defense Intelligence Agency (DIA) expanded the NIC's scale to create two scales of likeliness expressions ranging from *impossible* to *certain*

FIGURE 17.2 UK Defence Intelligence Uncertainty Yardstick

Qualitative Term	Associated Probability Range
Remote *or* highly unlikely	Less than 10%
Improbable *or* unlikely	15-20%
Realistic probability	25-50%
Probable *or* likely	55-70%
Highly probable *or* highly likely	75-85%
Almost certain	More than 90%

Source: Reproduced with permission of the UK government.

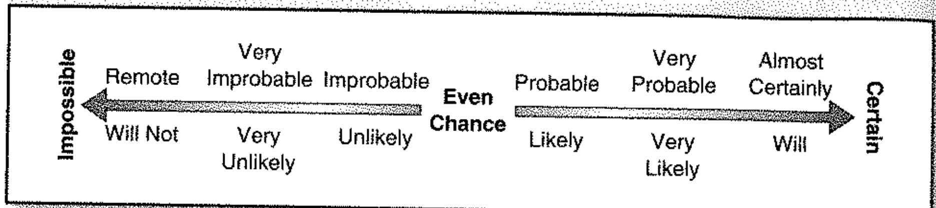
FIGURE 17.3 US National Intelligence Council Probability Scale

National Intelligence Council Probability Scale

Remote	Very Unlikely	Unlikely Chance	Even (Likely)	Probably	Very Likely	Almost Certainly

Source: Reproduced with permission of the US government.

FIGURE 17.4 US Defense Intelligence Agency Likeliness Expressions and Confidence Levels



Source: Reproduced with permission of the US government.

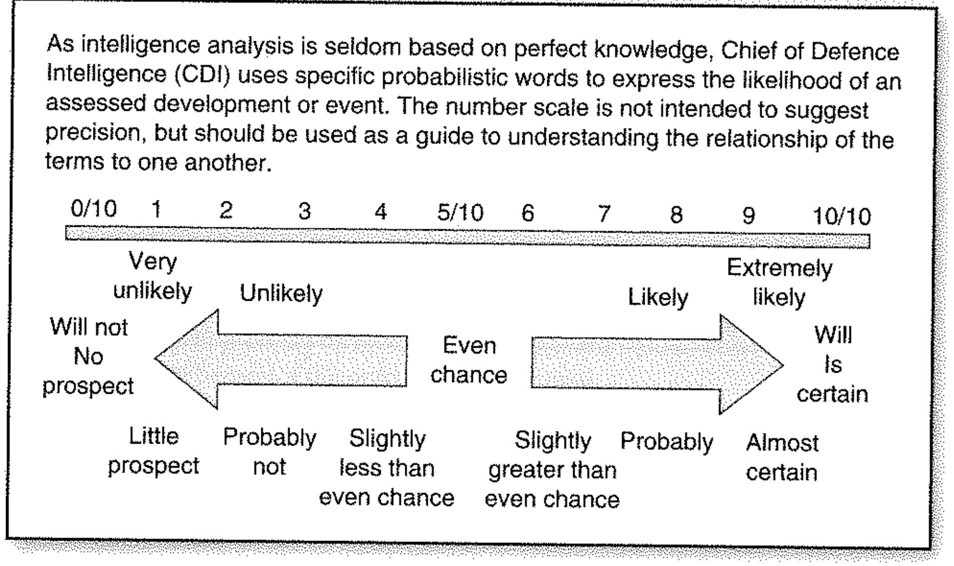
(see Figure 17.4). One scale focuses on likelihood and the other on probability. DIA also added a helpful caveat that, in the rare circumstance when likelihood cannot be assessed, analysts should use terms such as *may, could, might,* and *possibly.*[5] The word *could,* for example, would be used appropriately in an article describing the source of a cyber attack that could have been launched by a foreign country or by a lone hacker. The problem is that it is almost impossible to pin down the origin of most cyber attacks.

Canadian intelligence analysts have also wrestled with this problem. The Chief of Defence, Director General Intelligence Production, opted to array probabilistic terms on a scale from no chances in ten (*will not, no prospect*) to ten chances in ten (*will, is certain*), as shown in Figure 17.5.[6] The chart contains a caution that the numbers on the scale are not intended to suggest precision. The scale is best used as a guide to show the relationship of the terms to one another.

When using probabilistic language, analysts should be alert to two common traps often encountered in using verb forms.

1. Try to minimize use of modal verbs such as *could* or *might* because sentences that contain them usually convey little useful information to the reader.
2. Avoid using two probabilistic phrases in the same sentence—for example, by saying that "something *might* happen because the following

FIGURE 17.5 Canadian Probability Terms



Source: Reproduced with permission of the Canadian government.

conditions *are likely* to occur." The better phrasing is to state that "something *might* happen because the following conditions *are present*" or "something *will* happen *if* the following conditions occur." Similarly, a 50 percent chance of a 30 percent chance is nothing more or less than a 15 percent chance.

Analysts should also be alert to when they are mixing modal terms (i.e., *may, might, could* express the possibility of occurrence) with probabilistic terms (i.e., *probably, likely, unlikely* express the likelihood of occurrence).

- Using two modal terms in the same sentence usually presents no problem. For example, saying, "A political coalition *might* form because a hung parliament *is possible*" works because the cause (a hung parliament) and the effect (a political coalition) both have a *possibility* of occurring or not occurring (the choices are "yes/no" or, mathematically speaking, "1/0").
- Including a modal term and a probabilistic term in the same sentence can confuse meaning because an event that *is possible* ("yes" or "1") is no longer *impossible* ("no" or "0") so by definition has a probability greater than zero. The sentence "A political coalition *might* form because a hung parliament *is likely*" makes little sense because we have already stated the effect (a political coalition) *is possible* so the

probability or likelihood of the cause (a hung parliament) is irrelevant. Note that the coalition would result *because* of the hung parliament, but it is the effect of its *possibility* and not the *likelihood*.
- Modal terms chained to probabilistic terms, however, can be meaningful in some contexts. For example, if we know that the government is capable of authorizing defensive measures in the event of an impending attack, an analyst could meaningfully say, "Defensive measures *might* be taken because an attack *is likely*." In this case, we are certain about the *possibility* the government can take the measures, but we are uncertain about the *likelihood* it will do so.

One strategy for dealing with this environment of chronic imprecision when using probabilistic language is to substitute numbers for words. The military analysis community generally has a preference for employing percentages. Usually percentages of probability are conveyed in deciles, stating that an event has a 20, 30, or 60 percent chance of occurring. This approach helps the drafter avoid the well-established imprecision of narrative probability terms, but it often is faulted for conveying more precision than the evidence or the situation would warrant.

A good technique for assessing the soundness of a percentage probability judgment is to ask, "What is the probability percentage of this hypothesis being wrong?" If the probability percentage of the hypothesis being wrong and the probability percentage of the hypothesis being right do not add up to 100, then the analyst needs to rethink the assessment.[7]

Another solution is to present probability ranges by saying, for example, that an event has a 20 to 40 percent chance of occurring instead of making a point prediction. Such an approach is recommended only when the analyst has sufficient information to justify establishing a distinct upper boundary as well as a lower boundary for the estimate. However, simply distributing probabilities over a range (for example, by saying something is somewhere between 20 and 40 percent probable) is meaningless and should be treated—in every respect—the same as saying something has a 30 percent chance of occurring. To illustrate, a 50 percent chance of 90 percent and a 50 percent chance of 10 percent, or a uniform distribution between 10 to 90 percent, are all exactly the same as 50 percent, with no distinguishing features on any theoretical level. Providing a wide range, for example, of 20 to 70 percent is also more likely to confuse than to educate the client.

In presenting numerical probability estimates, analysts and clients must understand that analytic judgments cannot convey the same degree of mathematical precision as rolling dice or dealing cards from a deck. Nor can they provide the same degree of precision as a factory's estimate of the likely percentage of defective products after accumulating years of production data. This is especially true for analytic judgments that deal with one-time events or events that have no directly comparable precedent.

A strategy that one of the authors believes worked well in presenting judgments in National Intelligence Estimates he drafted is to substitute bettor's odds for percentages. For example, instead of saying something was almost certainly going to occur, the analyst would say the event had a 9-in-10 chance of occurring. Similarly, an event an analyst might assign a rough probability of 30 percent would be portrayed as having a 1-in-3 chance of succeeding.

Bettor's odds appear to be more effective in conveying probabilistic assessments because they convey an implicit risk calculation that is more likely to resonate with the decision makers who assess risks daily as part of their jobs. Moreover, most people are more accustomed to dealing with odds than with percentages, and they usually are interpreted more accurately than percentages. One downside is that users of bettor's odds can easily overstate low probabilities with statements like "only 1 chance in 1,000."

In his book *Calculated Risks,* Gerd Gigerenzer, a cognitive psychologist at the Max Planck Institute for Human Development in Berlin, explores how people miscalculate risk and uncertainty.[8] He makes a strong case for recasting conditional probabilities in terms of natural frequencies—simple counts of events—rather than the more abstract notions of percentages, odds, or probabilities. He argues that it is easier to comprehend the sentence, "Based on recent studies, eight out of every 1,000 women have breast cancer" than the statement, "The probability that a woman has breast cancer is 0.8 percent." In tests he conducted of German and American doctors presenting a problem involving conditional probabilities, an overwhelming percentage got the answer wrong when using percentages; when he replaced the percentages with natural frequencies, nearly all of them got the correct answer, or close to it.[9]

The best strategy for dealing with probabilistic judgments is to follow the probability statement (regardless of whether one uses a word, a percentage, or a set of odds) with the word *because* and a response to complete the sentence. For example, "We believe the event is *highly likely* to happen *because* two necessary conditions are present and a key driver is gaining strength." The remainder of the paragraph or the section should explore these three significant reasons in more detail.

By providing clients with explicit language laying out why a specific word or percentage was selected, they can make their own independent calculations of the probability of the event occurring. This approach also allows clients to track the accuracy of the assessment over time, monitoring the situation to see if the arguments for its being right are growing stronger or weaker.

LEVELS OF CONFIDENCE

The key to presenting levels of confidence is to state not just how confident you are as an analyst but why you are confident. The formula is similar to the one used to document probability statements. For example, "We judge with *high confidence* that the following will happen *because* we have two independent

sources stating this is the intent of the key decision makers." If, on the other hand, the analyst has low confidence in a particular source or judgment, then it is important to state what additional information or what additional events must transpire to increase the analyst's level of confidence. It is not a failure to declare low confidence when forecasting future possibilities that might result from the interplay of dynamic forces and factors.

As previously noted, the US Intelligence Community as a result of the 9/11 attacks and Iraq WMD estimate has mandated that analysts describe their level of confidence in judgments and assessments.[10] Most intelligence agencies use definitions of high, medium, and low levels of confidence, similar to those provided in Figure 17.6.[11] While the sets of definitions parallel each other closely, the NIC formulations focus more on an analyst's confidence in the estimative judgments provided in the paper. The definitions used by law enforcement organizations, including the Federal Bureau of Investigation (FBI) and Department of Homeland Security (DHS), usually are tied more to levels of confidence in the sourcing.[12]

FIGURE 17.6 US Intelligence Community Confidence Level Definitions

NATIONAL INTELLIGENCE COUNCIL (NIC)

Our assessments and estimates are supported by information that varies in scope, quality, and sourcing. Consequently, we ascribe *high, moderate,* or *low* levels of confidence to our assessments, as follows:

High Confidence generally indicates that our judgments are based on high-quality information, and/or that the nature of the issue makes it possible to render a solid judgment. A "high confidence" judgment is not a fact or a certainty, however, and such judgments still carry a risk of being wrong.

Moderate Confidence generally means that the information is credibly sourced and plausible but not of sufficient quality or corroborated sufficiently to warrant a higher level of confidence.

Low Confidence generally means that the information's credibility and/or plausibility is questionable, or that the information is too fragmented or poorly corroborated to make solid analytic inferences, or that we have significant concerns or problems with the sources.

FEDERAL BUREAU OF INVESTIGATION (FBI)

High Confidence: Direct or high-quality intelligence from multiple sources or from a single highly reliable source, such as high-quality imagery, human intelligence, or signals intelligence. High confidence generally indicates that

(Continued)

(Continued)

the FBI's judgments are based on high-quality information or that the nature of the issue makes it possible to render a solid judgment.

Medium Confidence: Indirect or derived intelligence from multiple sources or from a single reliable source. Medium confidence generally indicates that the information is interpreted in various ways, that the FBI has alternate views, or that the information is credible and plausible to render a solid judgment.

Low Confidence: Little or no information available, intelligence from untested sources, or for which there is little or no corroboration. Low confidence generally means that the information is scant, questionable, or very fragmented; that it is difficult to make solid analytic inferences; or that the FBI has significant concerns or problems with the sources.

DEPARTMENT OF HOMELAND SECURITY (DHS)

High Confidence generally indicates that judgments are based on high-quality information from multiple sources or from a single highly reliable source, and/or that the nature of the issue makes it possible to render a solid judgment.

Moderate Confidence generally means that the information is credibly sourced and plausible, but can be interpreted in various ways, or is not of sufficient quality or corroborated sufficiently to warrant a higher level of confidence.

Low Confidence generally means that the information's credibility and/or plausibility is questionable, the information is too fragmented or poorly corroborated to make solid analytic inferences, or that DHS and the FBI have significant concerns or problems with the sources.

Source: Reproduced with permission of the US government.

DIA stands out from most other agencies in that it has taken a more rigorous—and highly commendable—approach to defining confidence levels by establishing three distinct metrics (see Figure 17.4):

1. The strength of the knowledge base, reflected in part in the quality of the sources
2. The number and importance of key assumptions used to fill key information gaps
3. The strength of the underlying logic, measured in part by the use of analytic techniques

These metrics are then used to assess the analyst's level of confidence along three dimensions:

1. From uncorroborated to well-corroborated information
2. From many to minimal assumptions
3. From mostly weak to strong logical inferences[13]

We caution that repeatedly describing confidence levels in most paragraphs of an article or estimate can diminish the readability of a document. This problem can be overcome by including an overall summary of levels of confidence in a text box (see Chapter 8). In parts of the US Intelligence Community, these text boxes are referred to as *source summary statements*. They summarize the author's evaluation of the credibility of the sources and analytic judgments used in the article. If placed at the beginning of a document and read first, the source summary statement helps the reader more efficiently assess the significance of the information and judgments as they read the article.

Given the subjective nature of assessing levels of confidence, conveying levels of confidence graphically is often preferable. For example, key judgments can be portrayed in a text box or matrix and the level of confidence associated with each key judgment can be indicated with symbols, colors, or degrees of shading. The same approach works well when presenting a list of indicators or a list of key assumptions.

When information is conveyed in a matrix, a final column can be added on the right with the letters *H*, *M*, or *L* or three different icons to convey high, medium, or low levels of confidence for what is represented in each row. Alternatively, the cells in the matrix can be shaded with different colors. In this case, it is important to choose appropriate shades—for example, a deep red would be a poor choice to represent *low* given the intensity of the color. In our experience, various shades of purple or blue are the most effective, in part because they work for people who are color blind.

QUANTITATIVE DATA AND STATISTICS

When writing, reading, or reviewing papers that contain quantitative analysis, remember these ten rules of the road for presenting or interpreting quantitative data and statistics:[14]

1. Be cautious in making categorical claims and skeptical when you come across products that do so. We live in an uncertain world and rarely can something be actually "proven" or "disproven."
2. Openly acknowledge the uncertainty inherent in any study by listing up front the key assumptions that underpin the analysis. Papers that do so usually merit serious attention.

3. Pay attention to how data are collected and the context in which the research was conducted. Be wary of convenience sampling. Be more trustful when random sampling and double-blind testing are used. Samples should be large enough to justify conclusions and representative of the entire population.

4. Do not make the common mistake of comparing raw numbers without adjusting for expected baseline differences. For example, California obviously has more car accidents than Arizona because California has many more cars. A more reasonable comparison would be accidents per person in each state. Similarly, trends in dollars should always be reported with adjustments made for inflation.

5. Account for the fact that data often loses relevancy over time. Based on the experience of practitioners, a reasonable standard is that data used for forecasting should be no more than three to six months old, data used in medical research should be no more than two years old, and sources used to support national security analysis should be no more than five years old. Obviously, these time frames will vary depending on the circumstances, but as the world becomes increasingly complex we can expect these time frames to become ever shorter.

6. When presenting judgments based on percentages, ask if the contrapositive is also true; often it is not. For example, if *X* is 70 percent, is it true that *Not X* is 30 percent? This may not be the case if the remaining 30 percent constitutes *Y*, *Z*, and *K*. When dealing with circumstances where only *X* or *Not X* can be true, ask yourself the following question: "If *X* is 70 percent, am I equally comfortable with saying *Not X* is 30 percent?" If not, then the 70 percent estimate needs to be adjusted.

7. Determine or be explicit in stating whether the *average* is the *mean* (the sum of figures divided by the number of figures), the *mode* (the figure that appears the most frequently), or the *median* (the figure in the middle where half the figures are larger and half are smaller). In a normal distribution (the bell curve), the *mean, mode, and median* tend to be about the same. In nonstandard distributions, however, these values can vary widely given the same set of data.

8. Note if categories have been aggregated to present an unusually large number by connecting a stream of terms with the word *or*. For example, "Seventy percent of oversized men report being physically assaulted, denied employment, or insulted because of their size."

9. Understand that quantitative studies may not show *causation* but may show *correlation*. Two variables can be correlated in several different ways:

   - *X* may cause *Y*.
   - *Y* may cause *X*.
   - *X* and *Y* may affect each other.
   - *Z* (a totally different variable) may cause *X* and/or *Y*.

   For example, ice cream sales are correlated with drownings in many parks. Do ice cream sales cause drownings? Or do drownings cause ice cream sales? Or does a third factor (warm temperatures) create the conditions for both?

10. Know the meaning of statistical significance: that a result is unlikely due merely to chance. The conventional (and arbitrary) threshold for declaring statistical significance is a probability of less than 5 percent or a *p*-value less than 0.05. With large sample sizes, you often will see statistically significant results, especially if there is a relationship between the variables; small sample sizes often do not yield statistical significance. Statistical significance does not mean practical significance. To declare practical significance, you need to determine whether the size of the difference is meaningful.[15]

KEY TAKEAWAYS

- Both analysts and clients vary widely in the meanings they assign to probabilistic terms such as *likely* or *probably*.
- The best way to convey a level of likelihood is to follow the probabilistic word, percentage, or bettor's odd with the word *because* and a response to complete the sentence that includes a list of key factors that support the judgment.
- A good technique for assessing the soundness of a numeric probability judgment is to check to see if the percentage of a hypothesis being wrong and the percentage of it being right add to 100.
- The key to presenting levels of confidence is for analysts to state not just how confident they are as analysts but why they are confident.
- A source summary statement is a powerful tool for giving readers an overall sense of an analyst's level of confidence and the quality of the sources used to support the analysis before they start reading the paper. It also helps the analyst reduce visual clutter in the main document.
- Take care when presenting statistics and quantitative data and be even more skeptical when interpreting what you read.

CONSIDERING THE CASE STUDY

Review Case Study V, "Yemen: An Expanding Security Threat?" and draft a final paragraph or two presenting an analysis that answers the question, "Will Yemen pose a major threat to regional stability in the next five years?"

- How would you assess the chances that Yemen will become a major problem threatening stability in the region over the next five years? What words, percentages, or bettor's odds would you assign to this judgment?
- What is the rationale behind your choice of a word, a percentage, or bettor's odds in the answer you provided above? In other words, if you added the word *because* to the end of the preceding sentence, what reasons would you give to complete the sentence?
- What level of confidence do you have in the sources used to support the analysis on whether Yemen is likely to become a major problem threatening stability in the region?
- What level of confidence do you have in your overall assessment regarding Yemen's future stability?

NOTES

1. The information used in this chapter is adapted from Pherson Associates training materials (www.pherson.org).
2. Most of the exercises were conducted with US Intelligence Community analysts, although similar results were recorded in classes taught elsewhere in the US government and in the private sector. One hypothesis that needs to be explored is whether such differences are more prominent among English speakers as opposed to those speaking other languages.
3. "Understanding and Intelligence Support to Joint Operations," Joint Doctrine Publication 2-00 (3rd ed.), Ministry of Defence, August 2010, 3-23.
4. In the wake of the 9/11 attacks on the United States and the Iraq WMD fiasco, all National Intelligence Estimates published by the National Intelligence Council now include introductory material describing the NIC's methodology in providing probabilistic language and describing levels of confidence. The estimate *Iran: Nuclear Capabilities and Intentions* (November 2007), for example, includes a section on probabilistic language titled "National Intelligence Estimates and the NIE Process" as well as a text box titled "What We Mean When We Say: An Explanation of Estimative Language." The NIE and the introductory language can be accessed at http://www.dni.gov/files/documents/Newsroom/Reports%20and%20Pubs/20071203_release.pdf.
5. Defense Intelligence Agency, "What We Mean When We Say," chart of "Likeliness Expressions and Confidence Levels," May 18, 2010, used in courses Pherson Associates teaches for US Intelligence Community analysts.

6. Thompson, "Aide Memoire on Intelligence Analysis Tradecraft."
7. We are indebted to Richards J. Heuer Jr. for suggesting this commonsense technique for adding rigor to the use of probabilistic statements.
8. Gerd Gigerenzer, *Calculated Risks* (New York: Simon & Schuster, 2002).
9. Steven Strogatz, "Chances Are," *The New York Times*, April 25, 2010, http://opinionator.blogs.nytimes.com/2010/04/25/chances-are/.
10. Office of the Director of National Intelligence, "Intelligence Community Directive 206."
11. National Intelligence Council, "What We Mean When We Say: An Explanation of Estimative Language" as it appears following the Scope Note in the National Intelligence Estimate, *Iran: Nuclear Capabilities and Intentions*, November 2007, www.dni.gov/press_releases/20071203_release.pdf. The chart is used in courses Pherson Associates teaches for US Intelligence Community analysts.
12. The definitions of DHS and FBI levels of confidence are used in courses Pherson Associates teaches for US Intelligence Community analysts.
13. Defense Intelligence Agency, "What We Mean When We Say." This chart is used in courses Pherson Associates teaches for US Intelligence Community analysts.
14. These rules of the road were derived from multiple sources, including Pherson Associates training materials; Darrell Huff, *How to Lie With Statistics* (New York: W. W. Norton & Company, 1982); Niel A. Manson, "How to Lie With Statistics: Lessons for Intelligence Analysts," presentation to the Five Eyes Conference, University of Mississippi, March 17, 2015; and Jeffrey Sauro, "What Does Statistically Significant Mean?" Measuring U, October 21, 2014, http://www.measuringu.com/blog/statistically-significant.php.
15. Sauro, "What Does Statistically Significant Mean?"