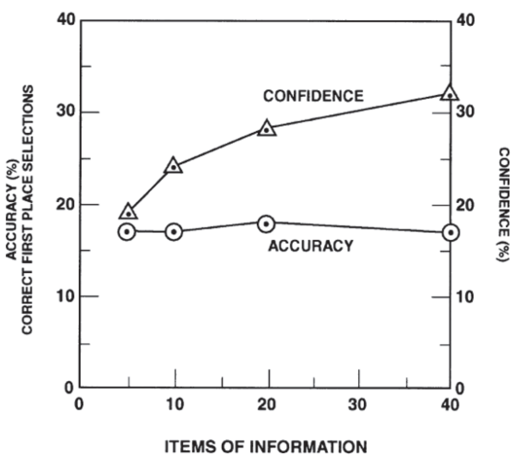
were asked to rank the top five horses in each race in order of expected finish. Each handicapper was given the data in increments of the 5, 10, 20 and 40 variables he had judged to be most useful. Thus, he predicted each race four times—once with each of the four different levels of information. For each prediction, each handicapper assigned a value from 0 to 100 percent to indicate degree of confidence in the accuracy of his prediction.

When the handicappers' predictions were compared with the actual outcomes of these 40 races, it was clear that average accuracy of predictions remained the same regardless of how much information the handicappers had available. Three of the handicappers actually showed less accuracy as the amount of information increased, two improved their accuracy, and three were unchanged. All, however, expressed steadily increasing confidence in their judgments as more information was received. This relationship between amount of information, accuracy of the handicappers' prediction of the first place winners, and the handicappers' confidence in their predictions is shown in Figure 5.

Figure 5



With only five items of information, the handicappers' confidence was well calibrated with their accuracy, but they became overconfident as additional information was received.

The same relationships among amount of information, accuracy, and analyst confidence have been confirmed by similar experiments in other fields.[55] In one experiment with clinical psychologists, a psychological case file was divided into four sections representing successive chronological periods in the life of a relatively normal individual. Thirty-two psychologists with varying levels of experience were asked to make judgments on the basis of this information. After reading each section of the case file, the psychologists answered 25 questions (for which there were known answers) about the personality of the subject of the file. As in other experiments, increasing information resulted in a strong rise in confidence but a negligible increase in accuracy.[56]

A series of experiments to examine the mental processes of medical doctors diagnosing illness found little relationship between thoroughness of data collection and accuracy of diagnosis. Medical students whose self-described research strategy stressed thorough collection of information (as opposed to formation and testing of hypotheses) were significantly below average in the accuracy of their diagnoses. It seems that the explicit formulation of hypotheses directs a more efficient and effective search for information.[57]

Modeling Expert Judgment

Another significant question concerns the extent to which analysts possess an accurate understanding of their own mental processes. How good is their insight into how they actually weight evidence in making judgments? For each situation to be analyzed, they have an implicit "mental model" consisting of beliefs and assumptions as to which vari-

55. For a list of references, see Lewis R. Goldberg, "Simple Models or Simple Processes? Some Research on Clinical Judgments," *American Psychologist*, 23 (1968), pp. 261-265.
56. Stuart Oskamp, "Overconfidence in Case-Study Judgments," *Journal of Consulting Psychology*, 29 (1965), pp. 261-265.
57. Arthur S. Elstein *et al., Medical Problem Solving: An Analysis of Clinical Reasoning* (Cambridge, MA and London: Harvard University Press, 1978), pp. 270 and 295.

ables are most important and how they are related to each other. If analysts have good insight into their own mental model, they should be able to identify and describe the variables they have considered most important in making judgments.

There is strong experimental evidence, however, that such self-insight is usually faulty. The expert perceives his or her own judgmental process, including the number of different kinds of information taken into account, as being considerably more complex than is in fact the case. Experts overestimate the importance of factors that have only a minor impact on their judgment and underestimate the extent to which their decisions are based on a few major variables. In short, people's mental models are simpler than they think, and the analyst is typically unaware not only of which variables *should* have the greatest influence, but also which variables *actually* are having the greatest influence.

All this has been demonstrated by experiments in which analysts were asked to make quantitative estimates concerning a relatively large number of cases in their area of expertise, with each case defined by a number of quantifiable factors. In one experiment, for example, stock market analysts were asked to predict long-term price appreciation for 50 securities, with each security being described in such terms as price/earnings ratio, corporate earnings growth trend, and dividend yield.[58] After completing this task, the analysts were asked to explain how they reached their conclusions, including how much weight they attached to each of the variables. They were instructed to be sufficiently explicit that another person going through the same information could apply the same judgmental rules and arrive at the same conclusions.

In order to compare this verbal rationalization with the judgmental policy reflected in the stock market analysts' actual decisions, multiple regression analysis or other similar statistical procedures can be used to develop a mathematical model of how each analyst actually weighed and combined information on the relevant variables.[59] There have been at least eight studies of this type in diverse fields,[60] including one involving

58. Paul Slovic, Dan Fleissner, and W. Scott Bauman, "Analyzing the Use of Information in Investment Decision Making: A Methodological Proposal," *The Journal of Business*, 45 (1972), pp. 283-301.
59. For a discussion of the methodology, see Slovic, Fleissner, and Bauman, op. cit.
60. For a list of references, see Paul Slovic and Sarah Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," *Organizational Behavior and Human Performance*, 6 (1971), p. 684.

These findings have broad relevance beyond the Intelligence Community. Analysis of information to gain a better understanding of current developments and to estimate future outcomes is an essential component of decisionmaking in any field. In fact, the psychological experiments that are most relevant have been conducted with experts in such diverse fields as medical and psychological diagnosis, stock market analysis, weather forecasting, and horserace handicapping. The experiments reflect basic human processes that affect analysis of any subject.

One may conduct experiments to demonstrate these phenomena in any field in which experts analyze a finite number of identifiable and classifiable kinds of information to make judgments or estimates that can subsequently be checked for accuracy. The stock market analyst, for example, commonly works with information concerning price-earnings ratios, profit margins, earnings per share, market volume, and resistance and support levels, and it is relatively easy to measure quantitatively the accuracy of the resulting predictions. By controlling the information made available to a group of experts and then checking the accuracy of judgments based on this information, it is possible to investigate how people use information to arrive at analytical judgments.

An Experiment: Betting on the Horses

A description of one such experiment serves to illustrate the procedure.[54] Eight experienced horserace handicappers were shown a list of 88 variables found on a typical past-performance chart—for example, the weight to be carried; the percentage of races in which horse finished first, second, or third during the previous year; the jockey's record; and the number of days since the horse's last race. Each handicapper was asked to identify, first, what he considered to be the five most important items of information—those he would wish to use to handicap a race if he were limited to only five items of information per horse. Each was then asked to select the 10, 20, and 40 most important variables he would use if limited to those levels of information.

At this point, the handicappers were given true data (sterilized so that horses and actual races could not be identified) for 40 past races and

54. Paul Slovic, "Behavioral Problems of Adhering to a Decision Policy," unpublished manuscript, 1973.

to the totality of material an analyst has available to work with in making a judgment.

Key findings from this research are:

- Once an experienced analyst has the minimum information necessary to make an informed judgment, obtaining additional information generally does not improve the accuracy of his or her estimates. Additional information does, however, lead the analyst to become more confident in the judgment, to the point of overconfidence.

- Experienced analysts have an imperfect understanding of what information they actually use in making judgments. They are unaware of the extent to which their judgments are determined by a few dominant factors, rather than by the systematic integration of all available information. Analysts actually use much less of the available information than they think they do.

As will be noted below, these experimental findings should not necessarily be accepted at face value. For example, circumstances exist in which additional information does contribute to more accurate analysis. However, there also are circumstances in which additional information—particularly contradictory information—decreases rather than increases an analyst's confidence.

To interpret the disturbing but not surprising findings from these experiments, it is necessary to consider four different types of information and discuss their relative value in contributing to the accuracy of analytical judgments. It is also helpful to distinguish analysis in which results are driven by the data from analysis that is driven by the conceptual framework employed to interpret the data.

Understanding the complex relationship between amount of information and accuracy of judgment has implications for both the management and conduct of intelligence analysis. Such an understanding suggests analytical procedures and management initiatives that may indeed contribute to more accurate analytical judgments. It also suggests that resources needed to attain a better understanding of the entire analytical process might profitably be diverted from some of the more costly intelligence collection programs.

prediction of future socioeconomic growth of underdeveloped nations.[61] The mathematical model based on the analyst's actual decisions is invariably a more accurate description of that analyst's decisionmaking than the analyst's own verbal description of how the judgments were made.

Although the existence of this phenomenon has been amply demonstrated, its causes are not well understood. The literature on these experiments contains only the following speculative explanation:

Possibly our feeling that we can take into account a host of different factors comes about because, although we remember that at some time or other we have attended to each of the different factors, we fail to notice that it is seldom more than one or two that we consider at any one time.[62]

When Does New Information Affect Our Judgment?

To evaluate the relevance and significance of these experimental findings in the context of intelligence analysts' experiences, it is necessary to distinguish four types of additional information that an analyst might receive:

- **Additional detail about variables already included in the analysis:** Much raw intelligence reporting falls into this category. One would not expect such supplementary information to affect the overall accuracy of the analyst's judgment, and it is readily understandable that further detail that is consistent with previous information increases the analyst's confidence. Analyses for which considerable depth of detail is available to support the conclusions tend to be more persuasive to their authors as well as to their readers.

- **Identification of additional variables:** Information on additional variables permits the analyst to take into account other factors that may affect the situation. This is the kind of additional information used in the horserace handicapper experi-

61. David A. Summers, J. Dale Taliaferro, and Donna J. Fletcher, "Subjective vs. Objective Description of Judgment Policy," *Psychonomic Science*, 18 (1970) pp. 249-250.
62. R. N. Shepard, "On Subjectively Optimum Selection Among Multiattribute Alternatives," in M. W. Shelly, II and G. L. Bryan, eds., *Human Judgments and Optimality* (New York: Wiley, 1964), p. 166.

ment. Other experiments have employed some combination of additional variables and additional detail on the same variables. The finding that judgments are based on a few critical variables rather than on the entire spectrum of evidence helps to explain why information on additional variables does not normally improve predictive accuracy. Occasionally, in situations when there are known gaps in an analyst's understanding, a single report concerning some new and previously unconsidered factor—for example, an authoritative report on a policy decision or planned coup d'etat—will have a major impact on the analyst's judgment. Such a report would fall into one of the next two categories of new information.

- **Information concerning the value attributed to variables already included in the analysis:** An example of such information would be the horserace handicapper learning that a horse he thought would carry 110 pounds will actually carry only 106. Current intelligence reporting tends to deal with this kind of information; for example, an analyst may learn that a dissident group is stronger than had been anticipated. New facts affect the accuracy of judgments when they deal with changes in variables that are critical to the estimates. Analysts' confidence in judgments based on such information is influenced by their confidence in the accuracy of the information as well as by the amount of information.

- **Information concerning which variables are most important and how they relate to each other:** Knowledge and assumptions as to which variables are most important and how they are interrelated comprise the mental model that tells the analyst how to analyze the data received. Explicit investigation of such relationships is one factor that distinguishes systematic research from current intelligence reporting and raw intelligence. In the context of the horserace handicapper experiment, for example, handicappers had to select which variables to include in their analysis. Is weight carried by a horse more, or less, important than several other variables that affect a horse's performance? Any information that affects this judgment influences how the handi-

Chapter 5

Do You Really Need More Information?

The difficulties associated with intelligence analysis are often attributed to the inadequacy of available information. Thus the US Intelligence Community invests heavily in improved intelligence collection systems while managers of analysis lament the comparatively small sums devoted to enhancing analytical resources, improving analytical methods, or gaining better understanding of the cognitive processes involved in making analytical judgments. This chapter questions the often-implicit assumption that lack of information is the principal obstacle to accurate intelligence judgments.[53]

* * * * * * * * * * * * * * * * * * *

Using experts in a variety of fields as test subjects, experimental psychologists have examined the relationship between the amount of information available to the experts, the accuracy of judgments they make based on this information, and the experts' confidence in the accuracy of these judgments. The word "information," as used in this context, refers to the totality of material an analyst has available to work with in making a judgment.

Using experts in a variety of fields as test subjects, experimental psychologists have examined the relationship between the amount of information available to the experts, the accuracy of judgments they make based on this information, and the experts' confidence in the accuracy of these judgments. The word "information," as used in this context, refers

53. This is an edited version of an article that appeared in *Studies in Intelligence*, Vol. 23, No. 1 (Spring 1979). That *Studies in Intelligence* article was later reprinted in H. Bradford Westerfield, ed., *Inside CIA's Private World: Declassified Articles from the Agency's Internal Journal*, 1955-1992 (New Haven: Yale University Press, 1995). A slightly different version was published in *The Bureaucrat*, Vol. 8, 1979, under the title "Improving Intelligence Analysis: Some Insights on Data, Concepts, and Management in the Intelligence Community." For this book, portions of the original article dealing with improving intelligence analysis have been moved to Chapter 14 on "Improving Intelligence Analysis."

capper analyzes the available data; that is, it affects his mental model.

The accuracy of an analyst's judgment depends upon both the accuracy of our mental model (the fourth type of information discussed above) and the accuracy of the values attributed to the key variables in the model (the third type of information discussed above). Additional detail on variables already in the analyst's mental model and information on other variables that do not in fact have a significant influence on our judgment (the first and second types of information) have a negligible impact on accuracy, but form the bulk of the raw material analysts work with. These kinds of information increase confidence because the conclusions seem to be supported by such a large body of data.

This discussion of types of new information is the basis for distinguishing two types of analysis- data-driven analysis and conceptually-driven analysis.

Data-Driven Analysis

In this type of analysis, accuracy depends primarily upon the accuracy and completeness of the available data. If one makes the reasonable assumption that the analytical model is correct and the further assumption that the analyst properly applies this model to the data, then the accuracy of the analytical judgment depends entirely upon the accuracy and completeness of the data.

Analyzing the combat readiness of a military division is an example of data-driven analysis. In analyzing combat readiness, the rules and procedures to be followed are relatively well established. The totality of these procedures comprises a mental model that influences perception of the intelligence collected on the unit and guides judgment concerning what information is important and how this information should be analyzed to arrive at judgments concerning readiness.

Most elements of the mental model can be made explicit so that other analysts may be taught to understand and follow the same analytical procedures and arrive at the same or similar results. There is broad, though not necessarily universal, agreement on what the appropriate model is. There are relatively objective standards for judging the quality of analysis, inasmuch as the conclusions follow logically from the application of the agreed-upon model to the available data.

Conceptually Driven Analysis

Conceptually driven analysis is at the opposite end of the spectrum from data-driven analysis. The questions to be answered do not have neat boundaries, and there are many unknowns. The number of potentially relevant variables and the diverse and imperfectly understood relationships among these variables involve the analyst in enormous complexity and uncertainty. There is little tested theory to inform the analyst concerning which of the myriad pieces of information are most important and how they should be combined to arrive at probabilistic judgments.

In the absence of any agreed-upon analytical schema, analysts are left to their own devices. They interpret information with the aid of mental models that are largely implicit rather than explicit. Assumptions concerning political forces and processes in the subject country may not be apparent even to the analyst. Such models are not representative of an analytical consensus. Other analysts examining the same data may well reach different conclusions, or reach the same conclusions but for different reasons. This analysis is conceptually driven, because the outcome depends at least as much upon the conceptual framework employed to analyze the data as it does upon the data itself.

To illustrate further the distinction between data-driven and conceptually driven analysis, it is useful to consider the function of the analyst responsible for current intelligence, especially current political intelligence as distinct from longer term research. The daily routine is driven by the incoming wire service news, embassy cables, and clandestine-source reporting from overseas that must be interpreted for dissemination to consumers throughout the Intelligence Community. Although current intelligence reporting is driven by incoming information, this is not what is meant by data-driven analysis. On the contrary, the current intelligence analyst's task is often extremely concept-driven. The analyst must provide immediate interpretation of the latest, often unexpected events. Apart from his or her store of background information, the analyst may have no data other than the initial, usually incomplete report. Under these circumstances, interpretation is based upon an implicit mental model of how and why events normally transpire in the country for which the analyst is responsible. Accuracy of judgment depends almost exclusively

Conclusion

There are many detailed assessments of intelligence failures, but few comparable descriptions of intelligence successes. In reviewing the literature on intelligence successes, Frank Stech found many examples of success but only three accounts that provide sufficient methodological details to shed light on the intellectual processes and methods that contributed to the successes. These dealt with successful American and British intelligence efforts during World War II to analyze German propaganda, predict German submarine movements, and estimate future capabilities and intentions of the German Air Force.[51]

Stech notes that in each of these highly successful efforts, the analysts employed procedures that "... facilitated the formulation and testing against each other of alternative hypothetical estimates of enemy intentions. Each of the three accounts stressed this pitting of competing hypotheses against the evidence."[52]

The simultaneous evaluation of multiple, competing hypotheses permits a more systematic and objective analysis than is possible when an analyst focuses on a single, most-likely explanation or estimate. The simultaneous evaluation of multiple, competing hypotheses entails far greater cognitive strain than examining a single, most-likely hypothesis. Retaining multiple hypotheses in working memory and noting how each item of evidence fits into each hypothesis add up to a formidable cognitive task. That is why this approach is seldom employed in intuitive analysis of complex issues. It can be accomplished, however, with the help of simple procedures described in Chapter 8, "Analysis of Competing Hypotheses."

51. Alexander George, *Propaganda Analysis: A Study of Inferences Made From Nazi Propaganda in World War II* (Evanston, IL: Row, Peterson, 1959); Patrick Beesly, *Very Special Intelligence: The Story of the Admiralty's Operational Intelligence Center 1939-1945* (London: Hamish Hamilton, 1977); and R. V. Jones, *Wizard War: British Scientific Intelligence 1939-1945* (New York: Coward, McCann & Geoghegan, 1978).
52. Frank J. Stech, Political and Military Intention Estimation: A Taxonometric Analysis, Final Report for Office of Naval Research (Bethesda, MD: MATHTECH, Inc., November 1979), p. 283.

of the 51 test subjects had the right answer the first time they thought they had discovered the rule.[50]

In the Wason experiment, the strategy of seeking confirming rather than disconfirming evidence was particularly misleading because the 2 - 4 - 6 sequence is consistent with such a large number of hypotheses. It was easy for test subjects to obtain confirmatory evidence for almost any hypothesis they tried to confirm. It is important to recognize that comparable situations, when evidence is consistent with several different hypotheses, are extremely common in intelligence analysis.

Consider lists of early warning indicators, for example. They are designed to be indicative of an impending attack. Very many of them, however, are also consistent with the hypothesis that military movements are a bluff to exert diplomatic pressure and that no military action will be forthcoming. When analysts seize upon only one of these hypotheses and seek evidence to confirm it, they will often be led astray.

The evidence available to the intelligence analyst is in one important sense different from the evidence available to test subjects asked to infer the number sequence rule. The intelligence analyst commonly deals with problems in which the evidence has only a probabilistic relationship to the hypotheses being considered. Thus it is seldom possible to eliminate any hypothesis entirely, because the most one can say is that a given hypothesis is unlikely given the nature of the evidence, not that it is impossible.

This weakens the conclusions that can be drawn from a strategy aimed at eliminating hypotheses, but it does not in any way justify a strategy aimed at confirming them.

Circumstances and insufficient data often preclude the application of rigorous scientific procedures in intelligence analysis—including, in particular, statistical methods for testing hypotheses. There is, however, certainly no reason why the basic conceptual strategy of looking for contrary evidence cannot be employed. An optimal analytical strategy requires that analysts search for information to disconfirm their favorite theories, not employ a satisficing strategy that permits acceptance of the first hypothesis that seems consistent with the evidence.

50. Harold M. Weiss and Patrick A. Knight, "The Utility of Humility: Self-Esteem, Information Search, and Problem-Solving Efficiency," *Organizational Behavior and Human Performance*, Vol. 25, No. 2 (April 1980), 216-223.

upon accuracy of the mental model, for there is little other basis for judgment.

It is necessary to consider how this mental model gets tested against reality, and how it can be changed to improve the accuracy of analytical judgment. Two things make it hard to change one's mental model. The first is the nature of human perception and information-processing. The second is the difficulty, in many fields, of learning what truly is an accurate model.

Partly because of the nature of human perception and information-processing, beliefs of all types tend to resist change. This is especially true of the implicit assumptions and supposedly self-evident truths that play an important role in forming mental models. Analysts are often surprised to learn that what are to them self-evident truths are by no means self-evident to others, or that self-evident truth at one point in time may be commonly regarded as uninformed assumption 10 years later.

Information that is consistent with an existing mind-set is perceived and processed easily and reinforces existing beliefs. Because the mind strives instinctively for consistency, information that is inconsistent with an existing mental image tends to be overlooked, perceived in a distorted manner, or rationalized to fit existing assumptions and beliefs.[63]

Learning to make better judgments through experience assumes systematic feedback on the accuracy of previous judgments and an ability to link the accuracy of a judgment with the particular configuration of variables that prompted an analyst to make that judgment. In practice, intelligence analysts get little systematic feedback, and even when they learn that an event they had foreseen has actually occurred or failed to occur, they typically do not know for certain whether this happened for the reasons they had foreseen. Thus, an analyst's personal experience may be a poor guide to revision of his or her mental mode.[64]

63. This refers, of course, to subconscious processes. No analyst will consciously distort information that does not fit his or her preconceived beliefs. Important aspects of the perception and processing of new information occur prior to and independently of any conscious direction, and the tendencies described here are largely the result of these subconscious or preconscious processes.

64. A similar point has been made in rebutting the belief in the accumulated wisdom of the classroom teacher. "It is actually very difficult for teachers to profit from experience. They almost never learn about their long-term successes or failures, and their short-term effects are not easily traced to the practices from which they presumably arose." B. F. Skinner, *The Technology of Teaching* (New York: Appleton-Century Crofts, 1968), pp. 112-113.

Mosaic Theory of Analysis

Understanding of the analytic process has been distorted by the mosaic metaphor commonly used to describe it. According to the mosaic theory of intelligence, small pieces of information are collected that, when put together like a mosaic or jigsaw puzzle, eventually enable analysts to perceive a clear picture of reality. The analogy suggests that accurate estimates depend primarily upon having all the pieces, that is, upon accurate and relatively complete information. It is important to collect and store the small pieces of information, as these are the raw material from which the picture is made; one never knows when it will be possible for an astute analyst to fit a piece into the puzzle. Part of the rationale for large technical intelligence collection systems is rooted in this mosaic theory.

Insights from cognitive psychology suggest that intelligence analysts do not work this way and that the most difficult analytical tasks cannot be approached in this manner. Analysts commonly find pieces that appear to fit many different pictures. Instead of a picture emerging from putting all the pieces together, analysts typically form a picture first and then select the pieces to fit. Accurate estimates depend at least as much upon the mental model used in forming the picture as upon the number of pieces of the puzzle that have been collected.

A more accurate analogy for describing how intelligence analysis should work is medical diagnosis. The doctor observes indicators (symptoms) of what is happening, uses his or her specialized knowledge of how the body works to develop hypotheses that might explain these observations, conducts tests to collect additional information to evaluate the hypotheses, then makes a diagnosis. This medical analogy focuses attention on the ability to identify and evaluate all plausible hypotheses. Collection is focused narrowly on information that will help to discriminate the relative probability of alternate hypothesis.

To the extent that this medical analogy is the more appropriate guide to understanding the analytical process, there are implications for the allocation of limited intelligence resources. While analysis and collection are both important, the medical analogy attributes more value to analysis and less to collection than the mosaic metaphor.

the same body of evidence may also be consistent with other hypotheses. A hypothesis may be disproved, however, by citing a single item of evidence that is incompatible with it.

P. C. Wason conducted a series of experiments to test the view that people generally seek confirming rather than disconfirming evidence.[49] The experimental design was based on the above point that the validity of a hypothesis can only be tested by seeking to disprove it rather than confirm it. Test subjects were given the three-number sequence, 2 - 4 - 6, and asked to discover the rule employed to generate this sequence. In order to do so, they were permitted to generate three-number sequences of their own and to ask the experimenter whether these conform to the rule. They were encouraged to generate and ask about as many sequences as they wished and were instructed to stop only when they believed they had discovered the rule.

There are, of course, many possible rules that might account for the sequence 2 - 4 - 6. The test subjects formulated tentative hypotheses such as any ascending sequence of even numbers, or any sequence separated by two digits. As expected, the test subjects generally took the incorrect approach of trying to confirm rather than eliminate such hypotheses. To test the hypothesis that the rule was any ascending sequence of even numbers, for example, they might ask if the sequence 8 - 10 - 14 conforms to the rule.

Readers who have followed the reasoning to this point will recognize that this hypothesis can never be proved by enumerating examples of ascending sequences of even numbers that are found to conform to the sought-for rule. One can only disprove the hypothesis by citing an ascending sequence of odd numbers and learning that this, too, conforms to the rule.

The correct rule was any three ascending numbers, either odd or even. Because of their strategy of seeking confirming evidence, only six of the 29 test subjects in Wason's experiment were correct the first time they thought they had discovered the rule. When this same experiment was repeated by a different researcher for a somewhat different purpose, none

49. Wason, ibid.

Evidence is diagnostic when it influences an analyst's judgment on the *relative* likelihood of the various hypotheses. If an item of evidence seems consistent with all the hypotheses, it may have no diagnostic value at all. It is a common experience to discover that most available evidence really is not very helpful, as it can be reconciled with all the hypotheses.

Failure To Reject Hypotheses

Scientific method is based on the principle of rejecting hypotheses, while tentatively accepting only those hypotheses that cannot be refuted. Intuitive analysis, by comparison, generally concentrates on confirming a hypothesis and commonly accords more weight to evidence supporting a hypothesis than to evidence that weakens it. Ideally, the reverse would be true. While analysts usually cannot apply the statistical procedures of scientific methodology to test their hypotheses, they can and should adopt the conceptual strategy of seeking to refute rather than confirm hypotheses.

There are two aspects to this problem: people do not naturally seek disconfirming evidence, and when such evidence is received it tends to be discounted. If there is any question about the former, consider how often people test their political and religious beliefs by reading newspapers and books representing an opposing viewpoint. Concerning the latter, we have noted in Chapter 2, "Perception: Why Can't We See What Is There to Be Seen?" the tendency to accommodate new information to existing images. This is easy to do if information supporting a hypothesis is accepted as valid, while information that weakens it is judged to be of questionable reliability or an unimportant anomaly. When information is processed in this manner, it is easy to "confirm" almost any hypothesis that one already believes to be true.

Apart from the psychological pitfalls involved in seeking confirmatory evidence, an important logical point also needs to be considered. The logical reasoning underlying the scientific method of rejecting hypotheses is that ". . . no confirming instance of a law is a verifying instance, but that any disconfirming instance is a falsifying instance."[48] In other words, a hypothesis can never be proved by the enumeration of even a large body of evidence consistent with that hypothesis, because

48. P. C. Wason, "On the Failure to Eliminate Hypotheses in a Conceptual Task," *The Quarterly Journal of Experimental Psychology*, Vol. XII, Part 3 (1960).

Conclusions

To the leaders and managers of intelligence who seek an improved intelligence product, these findings offer a reminder that this goal can be achieved by improving analysis as well as collection. There appear to be inherent practical limits on how much can be gained by efforts to improve collection. By contrast, an open and fertile field exists for imaginative efforts to improve analysis.

These efforts should focus on improving the mental models employed by analysts to interpret information and the analytical processes used to evaluate it. While this will be difficult to achieve, it is so critical to effective intelligence analysis that even small improvements could have large benefits. Specific recommendations are included the next three chapters and in Chapter 14, "Improving Intelligence Analysis."

If the hypothesis is incorrect, information may be lost that would suggest a new or modified hypothesis.

This difficulty can be overcome by the simultaneous consideration of multiple hypotheses. This approach is discussed in detail in Chapter 8. It has the advantage of focusing attention on those few items of evidence that have the greatest diagnostic value in distinguishing among the validity of competing hypotheses. Most evidence is consistent with several different hypotheses, and this fact is easily overlooked when analysts focus on only one hypothesis at a time—especially if their focus is on seeking to confirm rather than disprove what appears to be the most likely answer.

Failure To Generate Appropriate Hypotheses. If tentative hypotheses determine the criteria for searching for information and judging its relevance, it follows that one may overlook the proper answer if it is not encompassed within the several hypotheses being considered. Research on hypothesis generation suggests that performance on this task is woefully inadequate.[47] When faced with an analytical problem, people are either unable or simply do not take the time to identify the full range of potential answers. Analytical performance might be significantly enhanced by more deliberate attention to this stage of the analytical process. Analysts need to take more time to develop a full set of competing hypotheses, using all three of the previously discussed strategies—theory, situational logic, and comparison.

Failure To Consider Diagnosticity of Evidence. In the absence of a complete set of alternative hypotheses, it is not possible to evaluate the "diagnosticity" of evidence. Unfortunately, many analysts are unfamiliar with the concept of diagnosticity of evidence. It refers to the extent to which any item of evidence helps the analyst determine the relative likelihood of alternative hypotheses.

To illustrate, a high temperature may have great value in telling a doctor that a patient is sick, but relatively little value in determining which illness the patient is suffering from. Because a high temperature is consistent with so many possible hypotheses about a patient's illness, it has limited diagnostic value in determining which illness (hypothesis) is the more likely one.

47. Charles Gettys et al., *Hypothesis Generation: A Final Report on Three Years of Research.* Technical Report 15-10-80. University of Oklahoma, Decision Processes Laboratory, 1980.

The intelligence analyst has another tempting option not available to the policymaker: to avoid judgment by simply describing the current situation, identifying alternatives, and letting the intelligence consumer make the judgment about which alternative is most likely. Most of these strategies are not discussed here. The following paragraphs focus only on the one that seems most prevalent in intelligence analysis.

"Satisficing"

I would suggest, based on personal experience and discussions with analysts, that most analysis is conducted in a manner very similar to the satisficing mode (selecting the first identified alternative that appears "good enough").[46] The analyst identifies what appears to be the most likely hypothesis—that is, the tentative estimate, explanation, or description of the situation that appears most accurate. Data are collected and organized according to whether they support this tentative judgment, and the hypothesis is accepted if it seems to provide a reasonable fit to the data. The careful analyst will then make a quick review of other possible hypotheses and of evidence not accounted for by the preferred judgment to ensure that he or she has not overlooked some important consideration.

This approach has three weaknesses: the selective perception that results from focus on a single hypothesis, failure to generate a complete set of competing hypotheses, and a focus on evidence that confirms rather than disconfirms hypotheses. Each of these is discussed below.

Selective Perception. Tentative hypotheses serve a useful function in helping analysts select, organize, and manage information. They narrow the scope of the problem so that the analyst can focus efficiently on data that are most relevant and important. The hypotheses serve as organizing frameworks in working memory and thus facilitate retrieval of information from memory. In short, they are essential elements of the analytical process. But their functional utility also entails some cost, because a hypothesis functions as a perceptual filter. Analysts, like people in general, tend to see what they are looking for and to overlook that which is not specifically included in their search strategy. They tend to limit the processed information to that which is relevant to the current hypothesis.

46. The concept of "satisficing," of seeking a satisfactory rather than an optimal solution, was developed by Herbert A. Simon and is widely used in the literature on decision analysis.

Chapter 6

Keeping an Open Mind

Minds are like parachutes. They only function when they are open. After reviewing how and why thinking gets channeled into mental ruts, this chapter looks at mental tools to help analysts keep an open mind, question assumptions, see different perspectives, develop new ideas, and recognize when it is time to change their minds.

A new idea is the beginning, not the end, of the creative process. It must jump over many hurdles before being embraced as an organizational product or solution. The organizational climate plays a crucial role in determining whether new ideas bubble to the surface or are suppressed.

* * * * * * * * * * * * * * * * * * * *

Major intelligence failures are usually caused by failures of analysis, not failures of collection. Relevant information is discounted, misinterpreted, ignored, rejected, or overlooked because it fails to fit a prevailing mental model or mind-set.[65] The "signals" are lost in the "noise."[66] How can we ensure that analysts remain open to new experience and recognize

65. Christopher Brady, "Intelligence Failures: Plus Ca Change. . ." *Intelligence and National Security*, Vol. 8, No. 4 (October 1993). N. Cigar, "Iraq's Strategic Mindset and the Gulf War: Blueprint for Defeat," *The Journal of Strategic Studies*, Vol. 15, No. 1 (March 1992). J. J. Wirtz, *The Tet Offensive: Intelligence Failure in War* (New York, 1991). Ephraim Kam, *Surprise Attack* (Harvard University Press, 1988). Richard Betts, *Surprise Attack: Lessons for Defense Planning* (Brookings, 1982). Abraham Ben-Zvi, "The Study of Surprise Attacks," *British Journal of International Studies*, Vol. 5 (1979). *Iran: Evaluation of Intelligence Performance Prior to November 1978* (Staff Report, Subcommittee on Evaluation, Permanent Select Committee on Intelligence, US House of Representatives, January 1979). Richard Betts, "Analysis, War and Decision: Why Intelligence Failures Are Inevitable," *World Politics*, Vol. 31, No. 1 (October 1978). Richard W. Shryock, "The Intelligence Community Post-Mortem Program, 1973-1975," *Studies in Intelligence*, Vol. 21, No. 1 (Fall 1977). Avi Schlaim, "Failures in National Intelligence Estimates: The Case of the Yom Kippur War," *World Politics*, Vol. 28 (April 1976). Michael Handel, *Perception, Deception, and Surprise: The Case of the Yom Kippur War* (Jerusalem: Leonard Davis Institute of International Relations, Jerusalem Paper No. 19, 1976). Klaus Knorr, "Failures in National Intelligence Estimates: The Case of the Cuban Missiles," World Politics, Vol. 16 (1964).

66. Roberta Wohlstetter, *Pearl Harbor: Warning and Decision* (Stanford University Press, 1962). Roberta Wohlstetter, "Cuba and Pearl Harbor: Hindsight and Foresight," *Foreign Affairs*, Vol. 43, No. 4 (July 1965).

when long-held views or conventional wisdom need to be revised in response to a changing world?

Beliefs, assumptions, concepts, and information retrieved from memory form a mind-set or mental model that guides perception and processing of new information. The nature of the intelligence business forces us to deal with issues at an early stage when hard information is incomplete. If there were no gaps in the information on an issue or situation, and no ambiguity, it would not be an interesting intelligence problem. When information is lacking, analysts often have no choice but to lean heavily on prior beliefs and assumptions about how and why events normally transpire in a given country.

A mind-set is neither good nor bad. It is unavoidable. It is, in essence, a distillation of all that analysts think they know about a subject. It forms a lens through which they perceive the world, and once formed, it resists change.

Understanding Mental Ruts

Chapter 3 on memory suggested thinking of information in memory as somehow interconnected like a massive, multidimensional spider web. It is possible to connect any point within this web to any other point. When analysts connect the same points frequently, they form a path that makes it easier to take that route in the future. Once they start thinking along certain channels, they tend to continue thinking the same way and the path may become a rut. The path seems like the obvious and natural way to go. Information and concepts located near that path are readily available, so the same images keep coming up. Information not located near that path is less likely to come to mind.

Talking about breaking mind-sets, or creativity, or even just openness to new information is really talking about spinning new links and new paths through the web of memory. These are links among facts and concepts, or between schemata for organizing facts or concepts, that were not directly connected or only weakly connected before.

New ideas result from the association of old elements in new combinations. Previously remote elements of thought suddenly become as-

Strategies for Choice Among Hypotheses

A systematic analytical process requires selection among alternative hypotheses, and it is here that analytical practice often diverges significantly from the ideal and from the canons of scientific method. The ideal is to generate a full set of hypotheses, systematically evaluate each hypothesis, and then identify the hypothesis that provides the best fit to the data. Scientific method, for its part, requires that one seek to disprove hypotheses rather than confirm them.

In practice, other strategies are commonly employed. Alexander George has identified a number of less-than-optimal strategies for making decisions in the face of incomplete information and multiple, competing values and goals. While George conceived of these strategies as applicable to how decisionmakers choose among alternative policies, most also apply to how intelligence analysts might decide among alternative analytical hypotheses.

The relevant strategies George identified are:

- "Satisficing"—selecting the first identified alternative that appears "good enough" rather than examining all alternatives to determine which is "best."

- Incrementalism—focusing on a narrow range of alternatives representing marginal change, without considering the need for dramatic change from an existing position.

- Consensus—opting for the alternative that will elicit the greatest agreement and support. Simply telling the boss what he or she wants to hear is one version of this.

- Reasoning by analogy—choosing the alternative that appears most likely to avoid some previous error or to duplicate a previous success.

- Relying on a set of principles or maxims that distinguish a "good" from a "bad" alternative.[45]

45. Alexander George, *Presidential Decisionmaking in Foreign Policy: The Effective Use of Information and Advice* (Boulder, CO: Westview Press, 1980), Chapter 2.