Critical thoughts about critical thinking

Article *in* Theoretical Issues in Ergonomics Science · January 2010

DOI: 10.1080/1464536X.2011.564485

1 author:

Gary Klein
ShadowBox LLC & MacroCognition LLC

Critical thoughts about critical thinking

GARY KLEIN[1]

Klein Associates, A Division of Applied Research Associates, Inc. Fairborn, OH, USA

DRAFT manuscript

Submitted to
Theoretical Issues in Ergonomics Science

December 2009

[1] Email: gklein@ara.com

Abstract: For occupations such as intelligence analysis a great deal of effort has gone into developing critical thinking training and methods. These programs have a great deal of value, but they may also have drawbacks. This article discusses several problems with critical thinking programs and suggests that their emphasis on reducing mistakes needs to be balanced by an orientation to increase insights.

Key words: Critical thinking, insight, intelligence, mistakes

I. Introduction

Critical thinking is a central component of human intelligence, and is responsible for many achievements. Sternberg (1996) presented a tripartite model of the thinking skills responsible for successful intelligence: critical thinking, creative thinking, and practical thinking. For Sternberg, critical thinking involves analytical thinking such as analyzing, critiquing, judging, evaluating, comparing and contrasting, and is different from creative thinking and practical thinking. Halpern (2002) has described critical thinking as "…thinking that is purposeful, reasoned, and goal directed – the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions…" (p. 37). Halpern (2007) has claimed that critical thinking can be taught in schools. Her claim is based on the assumptions that there are clearly identifiable and definable critical thinking skills that students can be taught to apply.

Moore (2007) defined critical thinking as "…a deliberate meta-cognitive and cognitive act whereby a person reflects on the quality of the reasoning process simultaneously while reasoning to a conclusion. The thinker has two equally important goals: coming to a solution and improving the way she or he reasons (p. 8). Moore's definition highlights the metacognitive aspects, unlike Sternberg and Halpern, who are concerned with the analytical processes themselves. Because Moore's definition is broader it is more attractive. However, Moore's definition runs into the difficulties described by Nisbett and Wilson (1977) that people are very poor at introspecting about their own thinking processes; therefore, the agenda set forth by Moore may be impractical. Accordingly, I will rely on the description set forth by Sternberg and by Halpern.

Critical thinking seems to be a trainable skill. Examples of training in critical thinking include Rubinstein and Firstenberg (1987) and Woods (1987). Elder and Paul (2007) have published a critical thinking manual: "Analytical thinking: How to take thinking apart and what to look for when you do." Fischer, Spiker, Harris, and McPeters (2004) developed a program of computerized training in critical thinking for U.S. Army personnel. Several researchers (e.g., Kosonen & Winne, 1995) have found positive transfer from critical thinking training programs to work settings. Halpern points to the conclusion of the Thinking Skills Review Group (2005) that most studies show positive results from critical thinking training. Nisbett (1992) and his colleagues have also demonstrated positive outcomes for critical thinking training.

Not all of the results have been positive. Some researchers (e.g., Anderson & Soden, 2001; McMurray, Thompson, & Beisenherz, 1989) reported that critical thinking skills did not strongly transfer to a new domain.

The purpose of this paper is to explore some of the boundary conditions for critical thinking training. Studies that report positive effects may miss some of the limitations and drawbacks of critical thinking. It may strike some readers as odd to talk about drawbacks of critical thinking, but critical thinking programs may be implemented in isolation from the other dimensions of intelligence that Sternberg described such as creativity and practical reasoning. If a study only looks at the effects of a critical thinking program in reducing mistakes, it could miss any negative effects on functions such as creativity and practical reasoning.

In order to view critical thinking in a context, I examine the attempt to apply critical thinking standards in a practical domain – the field of intelligence analysis. The intelligence community (IC) cares deeply about the possibility of making erroneous or biased assessments and has applied critical thinking criteria to the way it judges intelligence products. According to Tam (2009), critical thinking programs have been introduced into a variety of schools run by different intelligence organizations. Critical thinking was added to the curriculum at the CIA's Sherman Kent School as of 2003 and to the training program at the Defense Intelligence Agency (DIA) in 2005. Similar critical thinking training has been implemented as a 40-hour program in the National Security Agency (NSA), and the Regional Joint Intelligence Training Facility in the United Kingdom, additional programs are also being offered. The IC programs teach intelligence analysts about checking assumptions, analyzing competing hypotheses, and so forth. This paper considers some of the drawbacks for the IC of taking such an approach, and then draws some implications about the use of critical thinking in general.

If errors in the IC are simply a result of failure to follow standards for keeping an open mind in reviewing data, considering all hypotheses, tracking assumptions, and being wary of judgment biases, then it is reasonable to direct more effort into developing and enforcing standards. However, I will argue that errors are not due to carelessness and failure to follow standards and that many of the standards currently in use may interfere with the way analysts think and may degrade their performance.

II. P = ↑ + ↓

Performance in domains such as intelligence analysis depends on increasing insights into what is going on, and reducing invalid conclusions (as shown in Figure 1). P = ↑ + ↓ means that performance depends on increasing insights and on reducing mistakes. Both of these activities are important. Analysts, who make accurate speculations but also become known for uncritical acceptance of flawed data, or for making erroneous arguments, will lose their credibility. On the other hand, a purely defensive strategy that documents all assumptions and evaluates all data sources but fails to notice connections and implications is not very helpful either. Avoiding mistakes is not the same as gaining insights. Increased amounts of rigor do not result in valuable discoveries. Yet senior intelligence officials admit that new analysts are often told that they need not worry about making incorrect judgments as long as they followed procedures and documented their sources, identified assumptions, and estimated uncertainty values. The products that analysts generate are evaluated for failure to adhere to standards; no one tracks whether the intelligence products contained valuable insights.

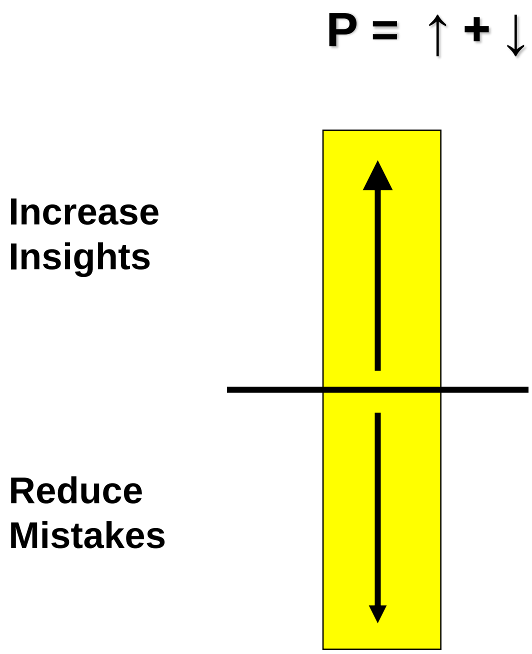

Figure 1. Performance depends on increasing insights and reducing mistakes.

Duggan (2007), in his book *Strategic Intuition*, discusses insight (the up arrow) as a process of noticing connections and relationships. Some people, because of their personal history and experiences and mindsets, are able to see connections that are invisible to others.

Within the intelligence community, performance evaluations are primarily associated with the down arrow, reducing mistakes, through the application of critical thinking practices. For example, Intelligence Community Directive Number 203, Analytic Standards, directs intelligence analysts to track assumptions, properly describe the quality and reliability of underlying sources, caveat and express uncertainties or confidence in analytic judgments, distinguish between underlying intelligence and analysts' assumptions and judgments, identify critical assumptions, consider alternative analysis where appropriate, use logical argumentation (e.g., maintain internal consistency, acknowledge supporting and contrary information), make accurate judgments and assessments, and engage in other practices that should reduce mistakes.

The two arrows, gaining insights and reducing mistakes, seem related to the System 1/System 2 formulation in cognitive psychology (Kahneman & Frederick, 2002; Kahneman & Klein, 2009). System 1 refers to rapid, automatic, associative thinking, such as pattern matching. System 2 refers to deliberate, logical, analytical thinking. Often, System 2 is viewed as a way to monitor the associations and impressions generated by System 1. Kahneman and Klein (2009) point out that both systems are necessary and that people will benefit from increasing their abilities in both directions. Klein (2009) has suggested that each system monitors the other. Not only does the critical thinking process of System 2 monitor the pattern-matching of System 1, but in addition the intuitive pattern matching of System 1 monitors whether the careful analyses of System 2 are plausible.

How do the two arrows relate to each other? One possibility is that reducing mistakes will also increase insights. I am not aware of any evidence that this is the case. A second possibility is that they are independent and unrelated.

The most disturbing possibility is that the actions taken to reduce mistakes can interfere with insights. In discussions with many experienced Intelligence Analysts, this third possibility seems to be the most likely, and it worries them. One highly regarded analyst explained that "In almost all useful products I've seen or done, there was some violation of tradecraft." In the intelligence community, the concept of "tradecraft" is often used as a catch-all term for systematic analytic techniques. In some uses, it is used as a synonym for "critical thinking," emphasizing valid arguments, careful vetting of sources, tracking of assumptions, ensuring logical conclusions, and so forth. Tradecraft emphasizes the down arrow, the reduction of mistakes. Reducing mistakes is critical and important, but the way it is implemented may pose difficulties.

How could the pursuit of the down arrow, the effort to reduce mistakes, interfere with insights? There are several ways this could happen.

i- **Anomalies.** The focus on tracking trends might lead analysts to overlook disruptions and breaks from tradition. Thus, during the Cuban Missile Crisis, John McCone, the head of the CIA, looked at the evidence of construction of missile defense systems around Cuba and judged that these had to be defending something – not a centralized HQ but something dispersed in different locations, and the most likely candidate was a set of ballistic missiles. On the other hand, Sherman Kent, a Senior Intelligence Analyst at the CIA, argued that in his study of Soviet decisions, such a risky and rash move was extremely unlikely. He had carefully reviewed historical data to understand the trends. Thus, his focus on historical trends led him to overlook the dynamics of the situation in Cuba.

ii- **Clutter.** The expansion of potential hypotheses may make it harder to intuit what is likely to be going on. The safest position, from the point of reducing mistakes, is to keep as many hypotheses in play as possible.

iii- **Distraction.** The process of tracking assumptions and uncertainties and checking the pedigree of sources, and making sure that arguments are not spoiled by flawed logic may interfere with insights that come from pattern-matching and associative reasoning. Wilson (2004) has shown that initial impressions reflect experience and avoid some of the bottlenecks of logical analysis, but Claxton (1999) has demonstrated that these initial impressions are disrupted when people start off by engaging in analysis. It is better to begin with initial impressions, realizing that they may be wrong.

iv- **Overshadowing.** Wilson (2004) has demonstrated that performance can suffer if people verbalize the cues they are using to make judgments because of verbal overshadowing – people focus on the cues they can articulate and ignore the tacit knowledge that is difficult or impossible to articulate. The emphasis on conscious thought required to list assumptions and identify areas of uncertainty may interfere with expertise that depends on nonverbal cues and tacit knowledge, according to unconscious-thought theory (Dijksterhuis & Nordgren, 2006). Dijksterhuis, Bos, van der Leij, and van Baaren (2009) have shown that experts who thought unconsciously outperformed experts who engaged in conscious thought and were better at applying diagnostic information. (The manipulations were arranged by presenting a problem and then having participants consciously think about the answer, the conscious thought condition, or shift their attention to an unrelated task, the unconscious thought condition.) This research, which is in line with the work of Wilson (2004), raises the possibility that instructions to engage in critical thinking – that is, to consciously deliberate – may prevent professionals from applying their expertise.

v- **Passivity.** The busywork of tracking assumptions and uncertainties may lead analysts to see their job in a passive way, as accountants rather than as detectives. An illustration of this passive attitude is the example cited earlier of junior analysts being told that the most important thing was that they follow tradecraft.

A study conducted by Rudolph and Raemer (2004) illustrates this last point. Rudolph studied 39 resident anesthesiologists in a life-like simulation setting to see how they interpreted evidence and diagnosed a problem. They were all taking a course in introductory crisis management in their anesthesia residency program. Based on their reactions to this scenario Jenny divided her subjects into four categories: Stalled, Fixated, Vagabonds, and Adaptive Problem Solvers.

Two of the anesthesiologists fit the “stalled” category. Neither of them could find any pattern that showed them how to proceed. They could not generate diagnoses and did not try different treatments; neither figured out the problem.

The 11 physicians categorized as fixated (including one Chief Resident at his hospital) usually jumped to the obvious diagnosis and rarely considered other hypotheses. None of these 11 anesthesiologists diagnosed the problem.

The 17 open-minded anesthesiologists fared no better. Jenny called their pattern “diagnostic vagabonding” because they wouldn’t commit to any diagnosis, but instead treated all possibilities as tentative. These physicians would consider each possibility but quickly jump to the others and never engage in a course of treatment that let them probe more deeply. None of them figured out the problem.

Last, we have the nine physicians who jumped to conclusions but tested those beliefs. Rudolph called them “adaptive problem solvers.” Like the fixated problem solvers, most of them identified the most likely cause right off the bat. But when their treatment didn’t work they turned to other diagnoses, testing and rejecting each. Their active exploration style let them use initial diagnoses as launching pads for conducting tests and treatments. This group nailed the problem; seven of these nine physicians discovered the problem. No physician in any of the other categories got there.

Rudolph expected that the anesthesiologists who jumped to a conclusion and held on to it would be unsuccessful, and she was right. None of them ever figured out the problem. Rudolph also expected that the anesthesiologists who kept an open mind while receiving the stream of information would be successful at making the right diagnosis, and here she was wrong. The ones who kept an open mind, absorbing data like sponges, also failed.

The only ones who succeeded had jumped to conclusions and tested them. They didn't fixate on their first explanation. Instead, they used that explanation to guide the tests they performed and the way they searched for new information. They exemplified the strategy of "strong ideas, weakly held."

Note that I am not advocating the abandonment of efforts to reduce mistakes or to engage in critical thinking; nor am I claiming that there is a zero-sum game. I am simply claiming that some of the procedures we enforce to reduce mistakes may interfere with gaining insights, and we need to be careful not to be too diligent in one direction or the other. The intelligence community has given some organizations the mission to ensure that analytical integrity and standards are implemented effectively. However, there are no corresponding organizations to advocate for the up arrow in Figure 1, which creates the possibility of an imbalance that may not serve the purposes of the intelligence community. Despite the current interest in analytical integrity, research investigations do not find that imposing analytical standards such as considering alternative hypotheses result in superior products, as measured by reduced errors or as increased insights (Cheikes, Brown, Lahner, & Alderman, 2004).

Part of the rationale for critical thinking programs is to reduce judgment biases by guiding people to think more systematically (Heuer, 1999). However, researchers have not demonstrated that debiasing techniques are practical or valuable (Yates, Veinott, & Patalano, 2003), or that it makes sense to try to stamp out judgment biases (e.g., Cohen, 1981; Cosmides & Tooby, 1996; Gigerenzer, 2005; Hertwig & Gigerenzer, 1999; Hoffman, 2005; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000; Klein, 2009) Lopes and Oden (1991) have reviewed the literature on judgment and decision biases, which form the basis for many critical thinking programs. They conclude that the concerns over judgment biases have been overstated. The studies have shown that under certain conditions, people will make systematic errors, but in natural settings these conditions do not arise very frequently. The heuristics that give rise to systematic errors are generally useful. Moreover, the recommendations for critical thinking that would replace or even augment human judgment are impractical in natural settings. They have three problems: robustness, generality, and practicality.

The so-called rational methods are not very *robust*. They are highly valid in that they lead to the right answers under ideal conditions. When facing difficult conditions such as noisy data, the strong methods become very brittle. The methods are not very *general*. For example, Bayesian techniques require a set of possible alternative hypotheses that are finite and specifiable in advance; however, in natural settings the hypotheses are usually not very easy to distinguish from each other. The informal reasoning methods people use do not lead to exact results, or require exact specification, making them more general. They are likely to be generally right, whereas the Bayesian methods are brittle under ambiguous conditions. The Bayesian methods simply have application requirements that are unrealistically strict. Finally, Lopes and Oden argue that the so-called rational methods are not *practical*. They are too laborious to apply, compared to heuristic methods that are much easier to use.

Perhaps it may be useful to consider a more balanced perspective on the work of intelligence analysts, based on Figure 1. Such a perspective would attempt to take advantage of both parts of the two-system model of cognition, intuition and analysis. The next section offers an initial description of such a perspective.

III. Dual Path Model of Data Interpretation

Based on a limited review of open-source material describing successes and failures of Intelligence Analysts and their customers, a dual-path model is proposed, as shown in Figure 2. The two paths are *insight* and *analysis*. They roughly correspond to the System 1/System 2 distinction discussed earlier. Both are operating in service of generating interpretations of situations. Figure 2 refers to "current interpretation," to reflect the temporary nature of any interpretation.

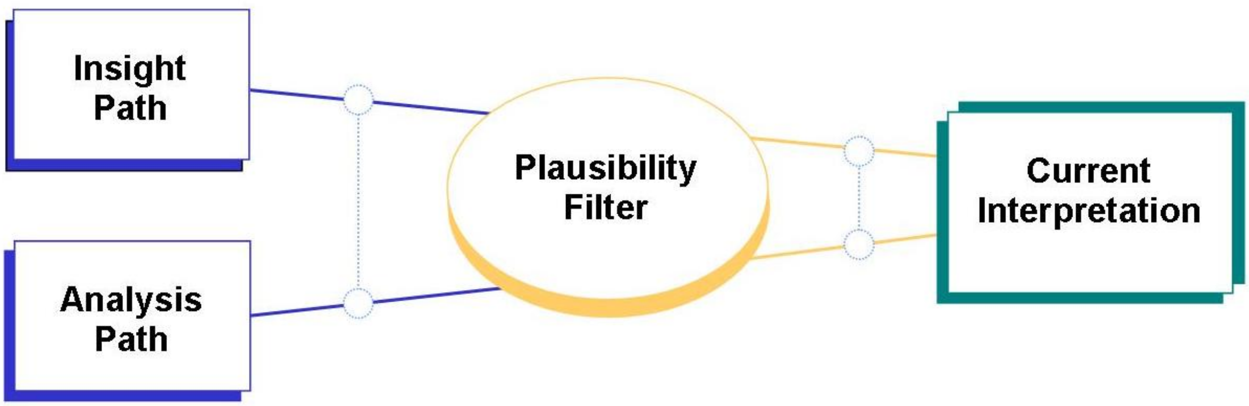

Figure 2. Dual path model of data interpretation.

The Insight Path

The Insight path stems from experience, from the patterns that experienced personnel such as Intelligence Analysts have learned, from their ability to see connections. It is associative and usually rapid. The military term *coup d'oeil* (eagle's glance) is sometimes used to describe how a skilled commander can look at a complex situation and rapidly recognize what to do. This path provides ballpark estimates and frames for more careful analysis. I will illustrate different aspects of the Insight path.

One example of the Insight path stems from the Battle of Taranto 11-12 November 1940 (Prange, 1981). The British used torpedo bombers launched from an aircraft carrier to attack and heavily damage an Italian naval fleet in the harbor of Taranto. The Italian battleship fleet lost half its strength in one night. The Japanese general Yamamoto reviewed the details of this battle and particularly noted that the torpedo attack worked despite the shallow harbor – only 40 feet deep. Previously, it had been believed that airborne torpedo attacks required at least 100 feet of water. Yamamoto first wrote about his plan to attack Pearl Harbor on 7 January, 1941.

The U.S. Chief of Naval Operations also took note of the Battle of Taranto; he signed a letter on 22 November, 1940 warning of a potential Japanese attack on the U.S. fleet at Pearl Harbor and advising the use of torpedo nets. His warning came before the first documented Japanese communication from Yamamoto, and was issued more than a year before the actual attack. He was not predicting an attack, merely noting its plausibility. His insight got lost during the subsequent year as military leaders misinterpreted Japanese intentions, but it still stands as an example of someone who made the connections and saw the implications.

The Insight path is responsible for detecting coincidence. It is not easy to spot coincidences because they are more than simple correlations, otherwise we would be detecting all kinds of connections. Coincidences are important because they reveal regularities that may have implications as we try to make sense of events. For example, one Intelligence Analyst described how he was reviewing transcripts of conversations from Serbian groups in Kosovo. He noticed repeated references to a Motorola type mobile telephone, and realized that this was being used as a primary C2 device. He noticed courteous expressions of welcome and departure at the beginning and the end of meetings, and realized that the meetings must have been taking place at people's homes. The coincidence enabled the analyst to make an inference that made it easier to locate the clandestine meetings. Similarly, in October 1962 an Intelligence Analyst noticed repeated references to the San Cristobal area in Cuba and added it to the target card used for U-2 aerial surveillance, enabling key discoveries about the Soviet build-up (Dobbs, 2008). Coincidence detection is not simply noticing similar elements because there are many ways events can be similar. Rather, coincidence reflects expertise in appreciating similarities that might be significant. It enables discovery of how the correspondence could be meaningful.

The Insight path has been responsible for a number of successes. In October 1962, the head of the CIA, John McCone, looked at the evidence regarding anti-missile defenses being constructed around Cuba and quickly speculated that they were intended to safeguard ballistic missiles that the Soviet Union must have been intending to install. Therefore, he directed his agency to use U-2 surveillance to keep track of Soviet progress. In contrast, Sherman Kent at the CIA downplayed the likelihood that the U.S.S.R. would be doing something so rash because it was not in keeping with their previous behaviors. Even afterwards, Sherman Kent argued that his analysis was sound – Khrushchev had behaved irrationally in taking the risks of sending missiles to Cuba. The incident reveals a limitation of careful analysis of historical trends; they are vulnerable to a discontinuous decision that doesn't follow the script.

Another example comes from the days leading up to Desert Storm, in 1991. An Intelligence Analyst noted the fact that the Iraqi government had commandeered approximately 10,000 trucks. The Kuwaitis would not have known about this action, so it was not intended as part of an Iraqi bluff. Therefore, he speculated that the Iraqis truly were preparing to invade Kuwait. He didn't base his speculation on any trend analysis, but on the connections he made about Saddam Hussein's preparation and intentions.

Continuing with Desert Storm, for the beginning of the ground phase of the war, the Marines had established a number of criteria for success. They were monitoring the time it took for the first units to get through the first part of the Iraqi defenses, and how long it took them to get through the second line, and how long it took their second wave to get through the Iraqi defenses. These were all traditional measures of progress. But then the word was received that thousands of Iraqi soldiers were surrendering. Gen. Norman Schwartzkopf was surprised by this news; he double-checked its accuracy. When it was verified, he moved up his timetable for the attack by 24 hours. The pattern he detected was of an enemy force that was collapsing and he wanted to take advantage of the momentum. The formal information requirements didn't request anyone to track the number of enemy who were surrendering, but when Schwartzkopf got this news, he immediately saw the implications.

Another example comes from the period prior to the Yom Kippur War. Israeli intelligence analysts noted that Egypt and Syria had evacuated their Soviet military advisors from Cairo and Damascus. They noted the departure of the Soviet fleet from Alexandria. They noted the movement of Egyptian antiaircraft missiles to the front lines. And in early October 1973 they began to issue warnings about the real possibility of an Arab attack. They had no clear evidence that the Egyptians were intending an attack, but they worried about the Israeli vulnerability because the events fit a pattern of preparation for an attack. But their insight was rejected because it was inconsistent with the analyses done by the head of Israeli intelligence (Bar-Joseph, 2005). He had determined that Syria wouldn't attack unless Egypt did, and that Egypt had no chance to win a war unless it had air superiority, or at least parity. His analyses blinded him to the connections between events.

When we review these examples, and add them to others that are well known (e.g., the FBI agents who wrote "the Phoenix memo" prior to the 9/11 attack describing incidents in which Middle Eastern Arabs had enrolled in U.S. flight schools but had not attended takeoff and landing training), we see sensemaking in context. There were no information requirements about flight schools or British naval attacks against Italian ships, or the movement of Soviet military advisors in Egypt and Syria in 1973. These data elements were only important in the context of the current situation. Each of them triggered the Insight path. None of them would have triggered the Analysis path because there none of them called forth historical records or trends.

The Insight path has also led to intelligence failures. The mistaken claim of Iraqi WMD is one example. One of the CIA intelligence analysts realized that the type of aluminum tubes Saddam Hussein was acquiring were perfectly suited for uranium enrichment, and concluded that Hussein was engaged in a clandestine program to construct nuclear weapons. It now appears that the tubes have other uses, such as missile launching. They are not ideal for this function, but in fact some countries, including Iraq, were using them in this way.

Another example of a mistaken insight is the FBI failure to identify the source of anthrax attacks. The FBI agents in charge of the investigation fixated on the wrong man, Dr. Steven Hatfill, and were unable to break loose from their conviction that he was the culprit. The FBI Director had to replace the leader of the anthrax investigation team with another agent in order to get a fresh look that resulted in the evidence against Dr. Bruce Ivins, who is now believed to be the perpetrator. These failures show how the Insight path can mislead us, and fixate us on a mistaken view. That is why insights have to be evaluated, using analytical methods. But it does not justify an emphasis on analysis that may interfere with insights.

The Analysis Path

The Analysis path is better studied and understood. It includes critical thinking and deliberate strategies. For example, Feltovich, Johnson, Moller, and Swanson (1984) described how physicians often generate a logical competitor set to capture the possible diseases that might be causing a given set of symptoms. The task of diagnosis is to use test results and run additional tests to rule members of this logical competitor set in or out. Such tests require us to generate a complete set of alternatives. Heuer's (1999) work on Alternative Competing Hypotheses is the way logical competitor sets are applied within the intelligence community. The use of Logical Competitor Sets provides a systematic way to perform diagnoses and make appraisals. It lends itself to various tools and techniques for gathering and sorting through evidence.

The Analysis path conforms well to recommendations for critical thinking, and to the tradecraft practices urged by groups supported better analytical integrity and standards. Zelik, Patterson, and Woods (2007) have discussed a variety of methods for injecting rigor into tradecraft. However, (Cheikes et al., 2004) demonstrated that ACH did not result in superior performance.

The Analysis path itself has resulted in some failures, such as the Sherman Kent interpretation of benign Soviet intentions in October 1962. One of the limitations of the Analysis path is that it involves viewing the world through a rear-view mirror, creating vulnerability to discontinuities in which the historical pattern is broken. Thus, turning again to the Cuban Missile Crisis, one reason that the U.S. intelligence analysts failed to identify some of the Cuban missile sites is that the security fences around these sites were either missing or incomplete. The analysts were used to seeing secure double fences – the Russians were diligent in protecting their missile sites in this way. However, due to time pressure, carelessness, or cultural differences, the missile sites in Cuba were not always ringed by the protective fences the Russians traditionally employed. As a result, the historical patterns were broken and the analysts relying on historical patterns failed to spot the sites (Dobbs, 2008).

Whereas the Analysis path seeks to expand possible hypotheses (see "A Tradecraft Primer: Structured analytic techniques for improving intelligence analysis", Central Intelligence Agency, 2005), the Insight path may limit the hypotheses that are considered. Damasio (1994) has noted that effective intuition lets us arrive at a solution without reasoning – not by assessing the entire field of possible options but by zeroing in on the appropriate ones. Damasio quotes Poincaré (1908) about how mathematical creation is not making new combination with entities already known, which would yield an infinite number of possibilities, most of them without any value. According to Poincare, creative intuitive insights depend both on seeing the possibility and on not making useless connections. "The sterile combinations do not even present themselves to the mind of the inventor." The process of avoiding useless pathways may employ a mechanism such as lateral inhibition, which in vision is useful for accentuating boundaries. With lateral inhibition, neurons in the visual system, when fired, not only record their stimulation but inhibit neighboring neurons. The result is a distorted but more definitive image. Perhaps something similar is happening to suppress alternatives in order to reduce uncertainty, increase confidence, and improve decisiveness. To the extent that this type of process is occurring, we need the Analysis path to overcome a tendency toward fixation. We also need the Insight path in order to prevent exponential explosion of possibilities.

Plausibility

As shown in Figure 2, the judgment of plausibility affects both the Insight path and the Analysis path. Without plausibility judgments, we would have to entertain all hypotheses indefinitely because there are usually some ways that a hypothesis might be accurate, explaining away contrary data. Plausibility judgments let us determine "No, this doesn't make sense." They let us prune the list of hypotheses. They let us direct our attention to the most likely/credible hypothesis. They let us make reasonable connections and discount unreasonable ones. For the Insight path, plausibility judgments determine our confidence in the beliefs and frames we are using. For the Analysis path, plausibility judgments keep us from entering into exponential explosions.

It is possible that the Insight and Analysis paths employ different types of plausibility judgments. There is very little research on how people make plausibility judgments. It is not by relying on Bayesian statistics, it is probably by trying to fit observed events into a frame and gauging the amount of cognitive strain. The more leftover events that the frame cannot handle, or the more the frame has to be jury-rigged to explain away inconsistencies, the greater the cognitive strain; the more reasonable the transitions, the more obvious the propensity for events to follow (be caused by) others, the greater the plausibility. Plausibility is probably related to the degree to which the narrative conforms to the criteria for a good story by tying together the intentions and actions of the agents with the observed events and the context of the situation. Koehler (1991) suggests that plausibility depends on the difficulty of mentally simulating a sequence of events.

A number of intelligence failures seem to involve plausibility judgments. The data were available, and individuals in the organization articulated the accurate hypothesis. But these early warnings were discarded because they varied too much from historical precedent or were deemed implausible. This was true for the Cuban Missile Crisis (Sherman Kent arguing against a Soviet introduction of missiles into Cuba), Desert Storm, 9/11, Pearl Harbor, and the Yom Kippur War. Perhaps this is why the use of tradecraft to emphasize the Analysis path seeks to minimize plausibility judgments. With hindsight from intelligence failures, we see that these judgments were wrong.

While we can find examples where plausibility judgments blinded people to the correct hypothesis, that doesn't mean we should stifle or seek to eliminate such judgments. To do so would commit the reasoning fallacy of backwards chaining: "People who are addicted to heroin drank milk when they were children. Therefore if we ban milk in schools we will reduce heroin addiction."

We see this backwards chaining fallacy in commercial aviation: "In all commercial aviation accidents the pilots have violated procedures. Therefore if we ensure that pilots follow procedures we will reduce accidents." But if you look at almost any flight you are likely to find some evidence of procedural violation. Even if unskilled pilots have more violations than skilled pilots and also have more accidents, that doesn't mean that reducing the procedural violations will increase safety. The real problem may be that inexperienced pilots are – inexperienced.

Relating this line of reasoning to the intelligence community, consider the following arguments: "When a forecast fails we can find evidence for mindsets. Therefore, if we eliminate mindsets, forecasts will improve." As with aviation, the correlation isn't there in the first place. You will find mindsets for successful forecasts as well as failures. Or substitute plausibility judgments for mindsets. The point is the same: the attempt to stamp out mindsets and plausibility judgments could backfire because the mindsets of the analysts are the way they apply their experience, and the plausibility judgments are the way that analysts prevent exponential explosions of hypotheses and focus their energies on the appropriate issues. Mindsets are the operation of mental models. The power of the mental models used by experienced personnel is that they provide plausibility judgments and guide attention. The weakness is that they can blind people to novelty. Mindlessly stamping out mindsets would prevent the experienced intelligence analysts from using their mental models to manage attention and notice anomalies.

IV. Conclusions

Using the domain of intelligence analysis as an example, I have identified a number of drawbacks to critical thinking programs: (a) Effective performance depends on gaining insights as well as reducing mistakes. Critical thinking methods are valuable for reducing mistakes, but may interfere with the process of gaining insights. If critical thinking methods are applied too enthusiastically they may interfere with insights and thereby reduce performance instead of improving it. (b) Critical thinking programs often include a metacognitive component to help people monitor and improve their thinking, but research has demonstrated that people are very poor at introspecting about their own thinking, which raises doubts about the value of metacognitive training; (c) the articulation of rationale the is typically required by critical thinking programs may reduce performance by interfering with unconscious thought processes that draw on tacit knowledge; (d) Critical thinking programs are often modeled on a standard of rationality that makes sense in controlled laboratory environments, but not in the complex situations that intelligence analysts and related specialties must operate. The critical thinking agenda is insufficiently robust, general or practical when applied to complex settings.

I am not advocating that critical thinking be abandoned; rather, I am pointing out that both insight and analyses are necessary for effective performance and for effective thinking. In line with the two-system model of cognition, I claim that both insight and analysis each has a role in monitoring the other, and in arriving at accurate and informed interpretations. The exaggeration of either process, either by over-emphasizing critical thinking, or by relying too heavily on intuition and eschewing critical thinking, is likely to result in poor performance.

The Insight path can be strengthened by working to build expertise more than to enforce standards. Expertise can be enhanced by using On-the-Job Training methods to transfer tacit knowledge. The intelligence community could try to develop ways to present weak signals to customers more effectively, so that they take alarms more seriously. They can try to shift the focus of current structured techniques–or developing new ones–to be more explicit and directive toward increasing and prompting insights. Gaining insights often involves fitting the same data into a new frame, or in discovering connections that others have missed.

How can the Insight path and the Analysis path be blended? This is a task for future researchers. However, it should be noted that Cohen, Freeman, and Thompson (1998) has proposed and demonstrated a recognition/metacognition approach aimed at achieving a productive synthesis. Also, the PreMortem method described by Klein (2007) also represents a blend of intuition and analysis. Hopefully, the proponents of critical thinking will be able to formulate more balanced strategies that can be safely applied in natural settings. The challenge is to design training, tools, and tactics for combining the Insight path and the Analysis path in a way that supports the strengths of each path and leads to more effective performance.

The emphasis on critical thinking needs to be matched by an emphasis on strengthening insights. This type of balanced perspective was on display when the American Philosophical Association initiated the Delphi project in 1988-89 to reach a consensus about critical thinking – what it is and what constitutes an effective critical thinker. They convened a panel of specialists in education, social sciences, and physical sciences. Facione (1990) presented the critical thinking cognitive skills they identified. These include practices for reducing mistakes, such as analyzing and assessing arguments, assessing claims, querying evidence, and justifying procedures. The list also includes practices regarding insights: decoding significance, examining ideas, conjecturing alternatives. In short, the panel formulated a set of practices that related both to insight and analysis.

The assertion portrayed in Figure 1 – that gaining insights is different from reducing mistakes – is relevant to many communities besides intelligence analysts. It is even relevant to the community of researchers. The tradecraft that researchers are taught in graduate school seems to focus on the downward arrow of Figure 1 – ways to reduce the chance of being wrong.

Imagine a continuum from statements that are very likely to be true, to statements that might well be wrong. And imagine a boundary line, separating the region of true statements from the region of possible errors (see Figure 3). In a risk-averse environment, researchers will be very reluctant to offer opinions that might be wrong. They will avoid the boundary line, moving as far to the accurate pole as possible (Klein, in press). That is fine – except that their statements and observations will carry very little information value. More desirable is a situation where researchers go in the opposite direction. They migrate to the boundary line, trying to make the most extreme observations that they believe they can defend. These are the insights, the risky predictions, and the disruptive opinions that can be so valuable.

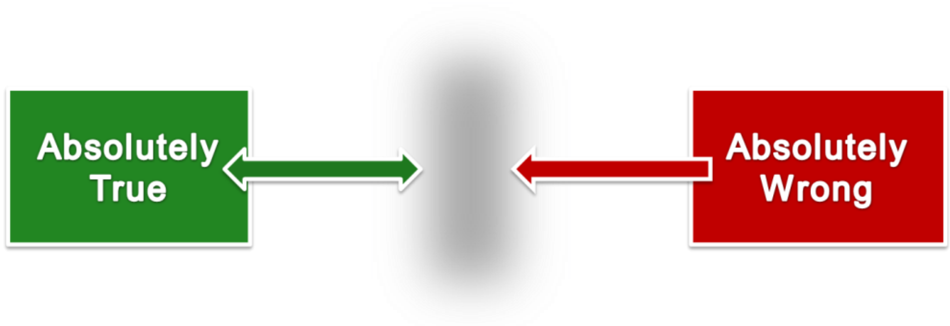

Figure 3. The trade-off of accuracy and information value.

Pressures that discourage errors can result in fewer insights, fewer discoveries. Critical thinking fuels the skepticism and doubts that are essential to good science. But doubt is insufficient to generate progress. The engine of science is curiosity, not caution. And curiosity is what drives the upward arrow towards the achievement of insight.

Acknowledgements: I would like to thank Neville Stanton, Julie Gore, Nicholas Sevdalis, Charles Mangio, and Bonnie Wilkinson for helpful comments on drafts of this paper.

V. References

Anderson, T., & Soden, R. (2001). Peer interaction and the learning of critical thinking. *Psychological Learning and Teaching*, *1*(1), 34-37.

Bar-Joseph, U. (2005). *The watchman fell asleep: The surprise of Yom Kippur and its sources*. Albany, NY: State University of New York Press.

Central Intelligence Agency. (2005). A tradecraft primer: Structured analytic techniques for improving intelligence analysis. *Tradecraft Review, 2*(2).

Cheikes, B. A., Brown, M. J., Lahner, P. E., & Alderman, L. (2004). *Confirmation bias in complex analyses* (Technical Report No. MTR 04B0000017). Bedford, MA: MITRE.

Claxton, G. (1999). *Hare brain, tortoise mind: How intelligence increases when you think less* (1st ed.). Hopewell, NJ: The Ecco Press.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *The Journal of Behavioral and Brain Sciences, 4*(3).

Cohen, M. S., Freeman, J. T., & Thompson, B. (1998). Critical thinking skills in tactical decision making: A model and a training method. In J. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 155-189). Washington, D.C.: APA Press.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*, 1-73.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York, NY: Penguin Books.

Dijksterhuis, A., Bos, M. W., van der Leij, A., & van Baaren, R. B. (2009). Predicting soccer matches after unconscious and conscious thought as a function of expertise. *Psychological Science, 20*, 1381-1387.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science., 1*, 95-109.

Dobbs, M. (2008). *One minute to midnight: Kennedy, Khrushchev, and Castro on the brink of nuclear war*. New York, NY: Knopf.

Duggan, W. (2007). *Strategic intuition: The creative spark in human achievement*. New York: Columbia University Press.

Elder, L., & Paul, R. (2007). *Thinker's guide to analytic thinking: How to take thinking apart and what to look for when you do*. Dillon Beach, CA: Foundation for Critical Thinking.

Facione, P. (1990). *Delphi research project, critical thinking: A state of expert consensus for purposes of educational assessment and instruction* (Research Report). Fullerton, CA: American Philosophical Association.

Feltovich, P. J., Johnson, P. E., Moller, J. H., & Swanson, D. B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W. J. Clancey & E. H. Shortliffe (Eds.), *Readings in medical artificial intelligence: The first decade* (pp. 275-319). Reading: Addison-Wesley.

Fischer, S. C., Spiker, A. A., Harris, D. H., & McPeters, E. R. (2004). *Critical thinking training for Army officers: Volume two: A model of critical thinking* (Technical Report). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gigerenzer, G. (2005). I think, therefore I err. *Social Research, 72*, 195-218.

Halpern, D. F. (2002). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Halpern, D. F. (2007). The nature and nurture of critical thinking. In R. J. Sternberg, H. L. R. III & D. F. Halpern (Eds.), *Critical thinking in psychology*. New York, NY: Cambridge University Press.

Hertwig, R., & Gigerenzer, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275-305.

Heuer, R. J., Jr. (1999). *Psychology of intelligence analysis*. Washington, D.C.: Center for the Study of Intelligence, Central Intelligence Agency.

Hoffman, R. (2005). *Biased about biases: The theory of the handicapped mind in the psychology of intelligence analysis.* Paper presented at the 48th Annual Meeting of the Human Factors and Ergonomics Society, Santa Monica, CA.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290*, 2261-2262.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics & Biases: The psychology of intuitive judgment* (pp. 49-81). New York, NY: Cambridge University Press.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise. *American Psychologist, 64*(6), 515-526.

Klein, G. (2007). Performing a project premortem. Harvard Business Review, September, 18-19.

Klein, G. (2009). *Streetlights and Shadows: Searching for the keys to adaptive decision making.* Cambridge, MA: MIT Press.

Klein, G. (in press). Looking for trouble. *Ethos*.

Koehler, D. (1991). Explanation, imagination, and confidence in judgement. *Psychological Bulletin, 110,* 499-519.

Kosonen, P., & Winne, P. H. (1995). Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology, 87*, 33-46.

Lopes, L. L., & Oden, G. C. (1991). The rationality of intelligence. In E. Eells & T. Maruszewski (Eds.), *Probability and rationality*. Amsterdam: Rodopi.

McMurray, M., Thompson, B., & Beisenherz, P. (1989). Identifying domain-specific aspects of critical thinking ability in solving problems in biology, *Annual meeting of the Southwest Educational Research Association*. Houston, TX.

Moore, D. T. (2007). *Critical thinking and intelligence analysis. Occasional Paper Number Fourteen*. Washington, D.C.: National Defense Intelligence College.

Nisbett, R. (1992). *Rules for reasoning*. Hillsdale, NJ: Lawrence Erlbaum & Associates, Inc.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 14*(1), 231-259.

Poincaré, H. (1908). Le Raisonnement mathématique, in Science et méthode (G. B. Halstead, Trans.). In B. Cheselin (Ed.), *The creative process*. Los Angeles, CA: Mentor Books/UCLA.

Prange, G. W. (1981). *At dawn we slept: The untold story of Pearl Harbor*. New York, NY: Penguin.

Rubinstein, M., & Firstenberg, I. (1987). Tools for thinking. In J. E. Stice (Ed.), *Developing critical thinking and problem solving abilities. New directions for teaching and learning, no. 30*. San Francisco, CA: Jossey- Bass.

Rudolph, J. W., & Raemer, D. B. (2004). Diagnostic problem solving during simulated crises in the OR. *Anesthesia and Analgesia, 98*(5S), S34.
Sternberg, R. J. (1996). *Successful intelligence: How practice and creative intelligence determine success in life*. New York, NY: Simon & Schuster.
Tam, C. K. (2009). *Behavioral and psychological considerations in intelligence analysis: A preliminary review of literature on critical thinking skills* (Interim Report). Mesa, AZ: Air Force Research Laboratory (AFRL).
Thinking Skills Review Group. (2005). *The impact of the implementation of thinking skills programmes and approaches on teachers* (Technical Report). Newcastle, UK: University of Newcastle, Learning and Research Center.
Wilson, T. (2004). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Harvard University Press.
Woods, D. (1987). How might I teach problem solving? In J. E. Stice (Ed.), *Developing critical thinking and problem solving abilities. New directions for teaching and learning, no. 30*. San Francisco, CA: Jossey-Bass.
Yates, J. F., Veinott, E. S., & Patalano, A. L. (2003). Hard decisions, bad decisions: On decision quality and decision aiding. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research*. New York, NY: Cambridge University Press.
Zelik, D., Patterson, E. S., & Woods, D. D. (2007). *Judging sufficiency: How professional intelligence analysts assess analytical rigor.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting, Baltimore, MD.