A Naturalistic Study of Insight

Gary Klein
Andrea Jarosz
MacroCognition LLC

Abstract: Although insight is often invoked as a phenomenon of problem solving and innovation, it has rarely been studied in a naturalistic fashion. The purpose of the study reported here was to learn more about insights as they occur in field settings as opposed to controlled laboratory conditions. The authors collected a set of 120 examples of insight taken from cognitive task analysis interviews, media accounts, and other sources and coded each incident using a set of 14 features. The results generated a descriptive model of insight that is different from the findings that emerge from research with puzzle problems. It posits multiple pathways for gaining insights. One pathway is triggered by detecting a contradiction. A second pathway is triggered by a need to break through an impasse. The third pathway gets triggered by seeing a connection.

Keywords: insight, impasse, mental model, discovery, frame, problem solving, innovation

Klein (2011) described two efforts that can be used to improve human performance. One of these is to reduce mistakes. Researchers and practitioners have developed a variety of methods for cutting down on error rates, including checklists, procedures, mnemonics, automatic error checkers, and even automatic error correctors. The second effort is to increase insights, which we define as discontinuous discoveries, that is, nonobvious inferences from the existing evidence. But we find few recommendations regarding ways to foster insights.

Fields such as intelligence analysis, for example, depend on both error mitigation and creative insight. No one wants intelligence analysts to apply sloppy reasoning methods but rather to be adept at critical thinking. However, eliminating mistakes is not the same thing as gaining insights. We need intelligence analysts who can gain insights into the motivations and the plans of adversaries and see the implications of subtle cues. The U.S. intelligence community has gone to

---

ADDRESS CORRESPONDENCE TO: Gary Klein, MacroCognition LLC, P.O. Box 533, Yellow Springs, OH 45387, gary@macrocognition.com.

Journal of Cognitive Engineering and Decision Making, Volume 5, Number 4, December 2011, pp. 335–351. DOI: 10.1177/1555343411427013.

great lengths to proceduralize critical thinking as a means of reducing mistakes (Heuer & Pherson, 2010). This emphasis on reducing mistakes stems in part from intelligence failures, for example, the mistaken analysis that the Iraqi government was probably building weapons of mass destruction prior to the 2003 invasion. To reduce the chances for such mistakes in the future, the director of national intelligence has a special office for ensuring analytical integrity. However, there is no corresponding office to promote the capability to gain insights—the insights that might have anticipated the 9/11 attacks. Much more attention is given to reducing mistakes than to encouraging insights.

The intelligence community is only a visible example of a trend that seems to exist in a wide variety of settings and organizations. Managers pay strict attention to following protocols and identifying departures from standard practices. They do not seem to give much attention to noticing, encouraging, and rewarding insights. More tools are available to follow procedures and reduce errors than to increase insights. There are no standards such as ISO 9000 and no techniques such as Six Sigma for promoting insights.

If the two efforts were complementary, then no harm would be done. Reducing mistakes would help people gain insights. Or, if the forces were unrelated, then the emphasis on reducing mistakes would not interfere with efforts to achieve insights. Unfortunately, in many situations the effort to reduce mistakes may potentially interfere with the achievement of insight by limiting time and resources and by directing attention toward precision and away from reflection. The effort at reducing mistakes—the documentation of sources and of areas of uncertainty, and the assignment of probabilities to assumptions—can get in the way of apprehending new patterns (Klein, 2011). Tracking historical trends too closely can mask disruptions that signal new trends. Critical thinking may encourage knowledge workers to view their jobs as not making mistakes rather than as gaining insights.

Given these practical considerations it might be valuable to empirically examine the second of the two efforts, the gaining of insights, to understand how insight works, what interferes with it, and how to support it.

Insight as Defined in the Psychology of Problem Solving

Psychologists have been studying insight for decades, starting with gestalt researchers such as Max Wertheimer (1945). The gestalt tradition viewed insight as a perceptual and conceptual reorganization, a sudden transformation. Gestalt researchers devised simple puzzle demonstrations that created impasses. The participants in these studies would quickly get stuck, unable to see how to arrive at a solution. Unlike straightforward problem-solving tasks, these puzzles could not be handled by steadfast effort. The puzzles were intended to lead the participant to an impasse. And then some of the participants would have an "aha!" experience and see how to reframe the issues and arrive at a solution. Sternberg and Davidson (1995) provided a comprehensive examination of a half century of this style of research. Most of the research that is done on insight involves these kinds of puzzles, college students as participants, and experimental tasks that last the duration of the typical college class.

In another tradition originating more in philosophy than psychology, scholars have studied famous cases, such as Archimedes's apocryphal "eureka" moment, Kekule's image of a snake as an idea of the structure of organic molecules, and Darwin's insight about a mechanism for evolution. However informative such case studies might be, most of what we have learned about insight has come from the controlled laboratory examination of how college students react when they reach the impasse created by puzzles. Weisberg (1995) presented a taxonomy of these puzzle problems. The taxonomy categories include brainteasers and riddles, geometrical problems (e.g., the area of a parallelogram, the nine-dot problem, the radiation problem), manipulative problems (the candle problem, the two-string problem), and mathematical problems.

Metcalfe and Wiebe (1987) found that participants' ratings of "warmth" (the nearness to a solution) increased as the participants got closer to finding a solution when working on noninsight problems, but the warmth ratings were flat for insight problems until the solution suddenly appeared. These data indicate that insight and noninsight problems have different features and that insight problems are marked by discontinuous discoveries. Gilhooly and Fioratou (2009) also obtained results differentiating insight and noninsight problems. Performance on the noninsight problems was linked to executive functions such as switching, whereas performance on the insight problems was not related to individual differences in executive functions, implying that insight problems are different from noninsight problems. They do not depend as heavily on System 2 processes involving abstract reasoning and hypothetical reasoning (Sloman, 1996; Stanovich & West, 2000).

We conducted the present study to learn more about insights as they occur in everyday situations. We were interested in sudden understanding but also in cases where people restructured their understanding even if these changes occurred gradually. Historical case studies of figures such as Darwin are informative, but there are not many examples like this, and they may not be representative of the environments in which most people live and work. The puzzle problem paradigm may also fail to capture some of the processes operating in natural settings. It may be that the puzzle problems do provide a useful framework for understanding everyday insights, but we will not know unless we conduct naturalistic investigations. This was the rationale for our project. We assembled a set of 120 incidents that were examples of insights and reviewed them to see if there were overarching themes and lessons to be learned.

Method

Incident Selection

We assembled a set of 120 incidents that were examples of insights. The criterion for inclusion was that the incident had to involve a person who made a radical shift in his or her mental model. This criterion distinguishes between *shifts* in a person's understanding and *elaborations* of the way a person understands a situation (Klein, Calderwood, & MacGregor, 1989). It distinguishes between episodes of sensemaking in which a person elaborates and preserves a frame from those in which a person reframes the situation (Klein, Phillips, Rall, & Peluso, 2007). The concept of a shift in understanding, a shift in a mental model, corresponds to the hypothesis (e.g., Gilhooly & Fioratou, 2009) that insight problems require people to restructure their initial problem representations, whereas noninsight problems can be solved by continued analyses without any restructuring.

The distinction between elaborations and shifts is illustrated by a case (Klein, 2005) in which a young nurse in a neonatal intensive care unit recorded the temperature of an infant in her charge. As the temperature dropped the nurse dutifully increased the warmth of the isolette. She was elaborating her understanding of the infant's current condition. A senior nurse stopped by, saw that the infant was still bleeding slightly from a heel stick made to get a blood sample, noticed that the infant's color seemed a bit mottled, read the chart that recorded the steadily falling temperature, and realized that the infant was coming down with sepsis—a dramatic shift in the understanding of the infant's condition.

Our study examined only shifts in which the person arrived at a different mental model that was more accurate, comprehensive, and useful. Sometimes the new mental model incorporated an additional causal factor that altered the existing factors in the mental model. Often the new mental model rejected parts or all of the previous one. The accounts of insights that we studied involved understanding what caused specific events, seeing new relationships between elements, or identifying new ways to accomplish an outcome. We did identify several false insights, cases in which the new mental model was incorrect, but did not examine them in the present study. These false insights did not meet the criterion of a shift to a different mental model that was more accurate.

The lead author collected instances of insight over a 6-month period. Of the incident accounts, 15 came from a review of cognitive task analysis interviews performed previously for other projects. Most of the incidents (45) came from books, particularly books describing examples of innovations and discoveries (e.g., Berkun, 2010; Hargadon, 2003; Johnson, 2010; Liedtka, Rosen, & Wiltbank, 2009; McGrath & MacMillan, 2009; Ogle, 2007; Perkins, 2000). Some (15) came from newspaper or magazine articles. Many (32) came from professional dialog and interaction including observations, discussions, and lectures. Another

TABLE 1. Domains Sampled in the Set of 120 Insights

Domain	Number Sampled in the Set of 120 Insights
Military	27
Invention	14
Science	12
Business and management	12
Investment	11
Medicine	9
Sports	5
Troubleshooting	5
Teaching	4
Design	4
Crime or detection	4
Firefighting	3
Politics	3
Miscellany	7

13 came from personal events and interactions involving the senior author as he encountered examples of insights formed by different people. (See Table 1 for a breakdown of the content of the 120 incidents.) Most were from the domains of military, financial investment, invention, science, business and management, and medicine.

Because of the method of collecting incidents, one concern is whether this corpus is representative. However, without an exploratory study such as this, it is not clear what the criteria for representativeness might be. We expect that future studies will be able to build on this one to identify better criteria for identifying incidents.

Each incident was briefly described in a textual description of one half to three pages, divided into three sections: background of the incident, critical events, and the nature of the insight.

Data Coding

The entire set of 120 incidents was coded independently by the two authors using 14 features. These features each had three or four possible values. Table 2 details the features used to assess insight incidents.

The set of 14 features emerged during successive reviews of the incidents as we searched for distinctions between the examples that might have implications for the nature of insight. Some of the features reflected theoretical issues, such as sudden versus gradual insights (no. 8), incubation (no. 9), and impasses (no. 12).

Most of the features evolved during the course of attempts to code the incidents. Sorting the 120 incidents into piles led us to distinguish insights that were triggered by contradictions (no. 2) as distinct from connections (no. 1). Then we

TABLE 2. Features Used to Assess Insight Incidents

Feature Name	Feature Description
1. Connections	The person made a connection between different data points or filled a gap (yes or no).
2. Contradictions	The person identified a contradiction in thinking (yes or no).
3. Explain away vs. explore	The person tried to explain away the contradiction or else explored it further.
4. Suspicious	The person had a suspicious or an open mind-set.
5. Understanding vs. action	The insight was about understanding, or understanding plus action.
6. Understanding vs. collaborative	The insight involved individual effort or collaborative efforts.
7. New data	The insight was triggered by new data versus a reorganization of thinking without any new data.
8. Sudden vs. gradual	The insight was sudden or gradual.
9. Incubation	There was vs. was not an incubation period.
10. Search	The insight was vs. was not about how to search for data.
11. Coincidence	The insight was vs. was not based on noticing coincidences.
12. Impasse	The person struggled with an impasse (yes or no).
13. Surprise	The person was vs. was not surprised.
14. Accidental	The insight was vs. was not accidental.

further differentiated some issues within the set of contradictions, whether the person was suspicious (no. 3) and whether the person accepted the contradiction rather than trying to explain it away (no. 4). We identified a small set of incidents that seemed to be about finding better ways to search for information (no. 10), which was a type of action. We distinguished a set of coincidences (no. 11) that seemed to be different from connections (no. 1) in that the coincidences were based on noticing repeated instances; noticing a coincidence was not itself an insight but led to the insight. We could distinguish insights that arose by accident (no. 14), usually involving coincidences or the detection of contradictions, from those that arose from deliberate, often desperate effort to solve a problem.

As can be seen, the set of features is very heterogeneous, rather than being conceptually equivalent. This study was designed as an initial exploratory investigation, and we chose to cast a wide net rather than prematurely settle on an overarching theoretical foundation for the selection of features.

Results

The interrater agreement was 78% for the initial, independent coding of instances. Cohen's kappa coefficient, a statistical measure of interrater agreement for qualitative items, was computed for each of the 14 features. The range was .17 to .66; by convention, kappa values less than .40 are considered poor

TABLE 3. Frequency Counts for the Different Features

Feature	A	B	Not Enough Information	No Rater Agreement
1. Connection	98	22		
2. Contradiction	45	74	1	
3. Explain away vs. explore	0	42	3	
4. Suspicious vs. open	26	13	3	3
5. Understanding and action	65	54		1
6. Individual vs. collaborative	82	35	1	2
7. New data	91	27		2
8. Sudden vs. gradual	54	42	11	13
9. Incubation	5	47	1	2
10. Search	16	104		
11. Coincidence	12	108		
12. Impasse	29	89	1	1
13. Surprise	109	9		2
14. Accidental	22	98		

agreement and kappa values between .40 and .75 are considered fair to good agreement (Banerjee, Capozzoli, McSweeney, & Sinha, 1999).

The raters obtained 98% agreement after comparing and discussing the features that they coded differently. The raters coded a sample, discussed their assignments, and repeated the process for the next sample. When the discussion led to a refinement of the description of a feature, the previous incidents were recoded to reflect that refinement. The kappa values for ratings following these calibration discussions were .89 or greater, with one exception. The rating for Feature 8 generated a kappa of .82. By convention, kappa values greater than .75 are considered excellent agreement. Table 3 details the frequency data. The explanations for columns A and B in Table 3 vary according to the feature and are described in the text below.

Feature 1 was about making a connection or filling a gap. For example, when Darwin read Malthus's speculations about the competition for scarce resources, he realized that resource competition was a force that could drive evolutionary changes. The majority of incidents (n = 98) were triggered when the person noticed a connection between different data elements; only 22 incidents (18%) did *not* show this feature.

Feature 2 was the discovery of a contradiction, which overlapped with seeing connections. When Harry Markopolos first looked at the kinds of financial returns that Bernard Madoff was reporting, he knew something was wrong because the strategy Madoff claimed to be using could not produce such regular and reliable rates of return year after year (Markopolos, 2010). Many of the incidents were initiated when a person encountered a contradiction. This occurred for 45 (38%) of 119 incidents (the raters selected "not enough information" for 1 of the incidents). The contradictions included seeing inconsistencies, seeing flaws in data that were widely accepted, or seeing flaws in the beliefs held by the person or by other people.

Feature 3 was a reaction to a contradiction, differentiating whether the person reacted to the contradiction by trying to explain it away or by exploring it further. This feature was not about the insight itself but examined the way the person reacted to the detection of a potentially important contradiction. For 42 of the 45 incidents (93%) that were initiated when a person spotted a contradiction, the person explored the ramifications of the contradiction rather than trying to explain it away. The Markopolos example illustrates this reaction. Once Markopolos spotted the contradiction, he devoted considerable energy to ferreting out Madoff's secrets. Three incidents did not have enough information to be coded for this feature.

Feature 4 contrasted suspicious versus open-minded stances when faced with contradictions. In 26 of these 45 cases that involved a contradiction, the person had a suspicious mind-set that enabled him or her to spot the contradiction or flaw. In only 13 of these contradiction insights did the person maintain an open mind. (For the remaining 6 incidents the raters either lacked enough information or did not reach agreement.) Thus, Markopolos was skeptical about Madoff's results even before he saw the record of results. However, seeing the regularity of the rate of return triggered the insight that Madoff was engaged in some sort of illegality.

Another example is the case of Meredith Whitney, a highly regarded Wall Street analyst who in 2008 was disinclined to give credence to rumors about problems at Bear Stearns. But then she decided to put on her skeptical lens and look at the publicly available data on Bear Stearns to see if she could make the case that the firm was in trouble. Now she started seeing problems she had previously ignored or explained away. As a result, she realized that Bear Stearns was facing a financial crisis (Cohan, 2009).

Feature 5 examined insights about action. Most instances, such as the Darwin, Markopolos, and Meredith Whitney examples, were just about gaining a better understanding (n = 65), but in many cases the insights suggested new affordances, new ways of acting (n = 54). The Wright brothers realized they could achieve sharp banked turns, instead of flat turns, by warping the wings of the airplane they were designing so that each wing had a different shape (Crouch, 1989). The raters did not reach agreement on one incident.

Feature 6 contrasted individual with collective insights. The majority of the cases were about individual effort (n = 82); a substantial number, 35 (30%), depended on collaborative effort such as the way Watson and Crick worked together to generate a double-helix model of DNA (Judson, 1996). Three cases were rated as lacking enough information or else did not reach agreement.

Feature 7 was whether the insight was triggered by new data. Most of the insights were triggered by new data (*n* = 91; 77%). The others were triggered by a reorganization of the person's mental model (*n* = 27; 23%). The raters failed to reach agreement on 2 of the incidents.

Feature 8 was sudden versus gradual insight. Most of the insights were sudden (*n* = 54; 56%), but many were gradual (*n* = 42; 44%). The raters were either unsure or unable to come to agreement on 24 incidents. This feature generated 13 failures to agree, even after discussion. None of the other features had more than 3 failures to agree. This feature also had the lowest kappa value, .82. The distinction was a difficult one to make because the published incident accounts were typically ambiguous about the emergence of the insight.

Feature 9 was about a period of incubation. Only 5 incidents explicitly noted a period of incubation; 65 incidents may have depended on incubation, but the data were insufficient to make a judgment. For 47 of the incidents no incubation period was possible. For example, in the Mann Gulch fire in Montana, a firefighter, Wagner Dodge, running uphill for his life ahead of a raging forest fire suddenly realized that he could set an escape fire ahead of himself, and he lived (Maclean, 1992). It seemed implausible that he spent any time incubating. The raters did not agree on 3 of the incidents.

Feature 10 considered whether the insight was about how to search for data. Thus, a paramedic needed to insert a breathing tube into the throat of a victim of a snowmobile accident but the victim's neck tissue was so badly mangled that the paramedic could not see where to insert the tube. Then he tried pressing on the victim's chest to force air out of the lungs and watched where bubbles appeared—that was the opening to the airway (Berlinger, 1996). The insight was not about where to insert the tube but rather about how act to gain a clue about the location of the opening. A small subset of cases, *n* = 16 (13%), were about better ways to search. The other 104 (87%) did not involve discovering new search strategies.

Feature 11 covered coincidences. Of the incidents, 12 (10%) depended on coincidences, such as John Snow noticing that victims of a cholera epidemic appeared to share a common water source (Johnson, 2006). The remaining 108 incidents (90%) did not.

Feature 12 was whether the incident involved an impasse. In only 29 of the incidents (25%) did the person reach an impasse, such as Wagner Dodge trying to outrun the forest fire. For 89 of the incidents (75%), no impasse was experienced. Two incidents were indeterminate in that there was either insufficient data or a failure of the raters to agree.

Feature 13 was whether the person was surprised by the insight. Most of the insights, 109 (92%), were accompanied by surprise, 9 (8%) were not, and the raters disagreed on 2. The prevalence of surprise reflects the discontinuous nature of insight. It may also reflect a tendency to recall and report instances of surprise because these make good stories. The nonsurprise examples of insights were those that did not contradict previous beliefs; they made the prior approach irrelevant. For example, in 1783 Napoleon, who was early in his career, pondered how to lift the British occupation of Toulon. He noted a small, lightly guarded fort in a hill overlooking Toulon. This fort was irrelevant to the previous French commander who was wrestling with the difficulties of attacking the British. Napoleon, however, was a specialist in the newly developed capability of light artillery. He worked out how the French forces could capture the small fort, bring up light artillery, and exploit this high ground advantage to force the British to withdraw (Duggan, 2007).

Feature 14 was whether or not the insight came about through an accident. In 22 (18%) of the incidents the insight was reached by accident, such as Fleming discovering penicillin (Perkins, 2000) or Page and Brin (the cofounders of Google) noticing that web searches included a number indicating links and then formulating a strategy for making use of those links (Duggan, 2007) or Jocelyn Bell Burnell noticing an anomaly while she was studying quasars and discovering the existence of pulsars (Colligan, 2009) or Michael Gottlieb wondering about a coincidence in the symptoms of several patients and discovering the outbreak of the AIDS epidemic (Shilts, 1987). The remaining 98 instances (82%) were not accidental—the person was deliberately working on the problem.

Discussion

A naturalistic study such as this is intended to open up dialog and an avenue for exploration of a cognitive function rather than providing disconfirming evidence or testing statistical hypotheses. Clearly, the study can be questioned on the basis of the set of incidents we selected. This selection was not random and may or may not be representative, whatever those terms might mean in this context. We reached for insight incidents that were ready at hand, rather than examples to support preexisting theories, but there is always a possibility that beliefs of which we were not aware might have governed the identification and selection of incidents. We wanted to cast a wide net that allowed us to explore different forms of insight.

Some insights were spurred by impasses, but most were not. Virtually all insights involved a change in understanding (this was one of our selection criteria), but many also involved action—a realization of how to make things happen. The new understanding was coupled with a discovery of new affordances.

Many insights involved seeing connections, but a surprising number of insights were triggered by inconsistencies and contradictions. The insights that were triggered by contradictions seemed to depend on the person taking the anomalous data point seriously rather than attempting to explain it away. Klein et al. (2007; Klein, Moon, & Hoffman, 2006) described a Data/Frame model of sensemaking that contrasted two reactions when people confronted evidence that questioned a frame that was explaining the dynamics of a situation. People could try to preserve the frame by making small elaborations to it, or they could take the evidence seriously and try to reframe their understanding of the situation. The first reaction seems to be more common. People use knowledge shields to preserve their frames (Chinn & Brewer, 1993; Feltovich, Coulson, & Spiro, 2001). The second reaction, reframing the situation, restructuring the way the situation is represented, is the one that leads to insights.

The data reveal and clarify some limitations in the standard laboratory paradigm for studying insight. This paradigm has been valuable for helping researchers understand many aspects of insight. However, the paradigm involves puzzles that force the person to experience an impasse. The impasse is created by the design of tasks and materials that lead the majority of participants to make inappropriate assumptions. The insight is triggered when the person realizes that one or more of these assumptions is unnecessary for the rules of the task (Smith, 1995). Such conditions do not fit the majority of the incidents that we studied, nor do the additional restrictions of having a specific starting point and ending point for gaining the insight, working on a task that is irrelevant to the person's current concerns, and eliminating the role of expertise by using artificial tasks. In the majority of the incidents we studied, the person's prior experience was critical for gaining the insight. We appreciate the advantages of the puzzle problem paradigm for enabling a wide variety of manipulations and even for use in neuroimaging studies (e.g., Bowden, Jung-Beeman, Fleck, & Kounios, 2005). Our suggestion is that this work could be described as studying the "aha!" experience rather than exploring the full variety of forms of insight.

The notion of insight as a sudden phenomenon may have to be reexamined because so many of the incidents in our sample were gradual (44%). We found that sudden insights were most likely to be triggered by new data, but the total context of the problem stretched considerably before and beyond the time frame for the appearance or discovery of new data. Gradual insights were more likely to be tied to long-standing hunches, as noted by Johnson (2010) in his examination of innovations. The hallmark of insight has been the "aha!" experience, so we were surprised by the finding that so many insights in our sample were gradual rather than sudden. The "aha!" experience (e.g., Topolinski & Reber, 2010) may actually be different from insight. We suggest that it is an epiphenomenon that accompanies some but not all insights. For example, Metcalfe and Wiebe (1987) have shown that insights, unlike routine problem solving, occur without warning. We believe that the Metcalfe and Wiebe findings stem from their use of the impasse paradigm. The "aha!" experience fits tasks such as ambiguous dot patterns that suddenly get recognized or puzzles that rely on participants suddenly realizing that they have been holding flawed assumptions. Some insight researchers (e.g., Ippolito & Tweney, 1995; Weisberg, 1995) do not view "aha!" as a criterion for insight.

How might gradual insights emerge? We found several routes. One was the detection of coincidences. When Gottlieb and his colleagues encountered their first AIDS patient, they were puzzled. By the fifth patient, they reported their findings on a new, mysterious, and terrifying disease. Someplace in between these five patients they spotted a repetition of symptoms that formed the cluster of their finding. Another route to gradual insight was a deliberate search, such as the one that Meredith Whitney conducted to see if Bear Stearns was in financial difficulty. As she uncovered more and more evidence, she came to realize that the firm was in deep trouble. A third route to gradual insights was an incremental process, as in the development of printing (Man, 2002) or the mass production process used by Ford Motor (Hargadon, 2003). The revolutionary technologies did not have any single "aha!" instant.

Our results challenge other common beliefs about insight. The process of incubation may be valuable for forming insights, but it does not seem necessary. The benefits of an open mind cannot be denied, but there may also be value in conducting a more focused investigation, particularly when the investigation challenges some widely accepted assumptions. The Meredith Whitney and Markopolos examples showed how insights can be driven by skeptical investigators. Whitney described how she deliberately looked at the Bear Stearns data through a skeptical lens.

Our findings were also inconsistent with the belief that expertise gets in the way of insight. Studies of mental set, or Einstellung, seem to find that when people get more experience they become less likely to see novel strategies. The Luchins and Luchins (1959) water jar experiment showed that giving participants repeated trials in which the same formula was successful in generating a solution would reduce their chances of finding a more efficient solution later on. However, the examples we studied did not provide nearly identical conditions for trial after trial. Most of the puzzles studied with impasse paradigms try to use participants' experience against them so that they will make dysfunctional assumptions. In contrast, we estimated that experience was necessary for two thirds of the insights in our sample of naturally occurring incidents.

Therefore, the topic of insight overlaps with the study of expert problem solving. Several of our examples were the types of medical diagnoses and scientific discoveries that have stimulated research into expert problem solving. A third of the cases in our sample did not depend on expertise, and several aspects of expert problem-solving strategies are not related to discontinuous discoveries, so the field of expert problem solving is distinct from the field of insight. Nevertheless, it might be useful for future investigations to explore the overlap in greater detail.

Based on our study, we define insight as a discontinuous discovery, a nonobvious revision to a person's mental model of a dynamic system, resulting in a new set of beliefs that are more accurate, comprehensive, and useful. Insights are discontinuous discoveries in that they do not directly follow from the data available prior to the insight. The person gaining the insight shifts his or her mental model; either new data or a combination of data or a finding of a contradiction leads the person into a conceptual territory that was new in some aspects. To define mental model, we follow the account of Doyle and Ford (1998) of a set of beliefs about the nature of phenomenon and how it works, as well as the definition and methodology laid out by Klein and Hoffman (2008).

Insight generally requires the person to abandon one or more beliefs. Klein et al. (1989) distinguished between shifts and elaborations of the way a person understands a situation. We view insight as the former, a shift in the way a person thinks, acts, sees, desires, and feels. It is a form of sensemaking that occurs when a person restructures his or her understanding.

Klein et al. (2007) described sensemaking as a process of using frames to organize and integrate data elements and beliefs. Mental models and stories are examples of frames. Sensemaking involves building, elaborating, preserving, questioning, revising, and replacing frames. Insights would fall into the last two categories, revising and replacing frames. Klein et al. also suggested that sensemaking was directed at the three to four most important data elements or beliefs—these were the anchors for the frame. From this perspective, we hypothesize that when a person revises a frame or constructs a new frame, what is happening is that these three to four anchors are being modified. Some of them are being altered or replaced, or a new anchor is being introduced into the mix. Figure 1 shows three pathways to insightful revisions in a person's mental model.

One pathway gets initiated when a person encounters a contradiction. Instead of trying to explain away the contradiction by rejecting the weakest anchor or finding some compromise, the person accepts the weakest anchor and revises the others. Meredith Whitney's first reaction on hearing that the investment firm Bear Stearns was in trouble was to dismiss the rumors. They contradicted all that she knew about the health of the company. But then she decided to see what would happen if she took the rumor seriously and arrived at a different belief—that Bear Stearns was unlikely to survive.

Something similar happened when Harry Markopolos started to look into Bernie Madoff's investment successes. Markopolos was surprised by the regularity of profits; Madoff seemed to make money every month regardless of what the stock market did. Yet Madoff was a highly respected figure in the financial community. So Markopolos had run into a contradiction. He followed the unlikely anchor, that Madoff was behaving illegally, and eventually uncovered the evidence for Madoff's Ponzi scheme.

A second pathway gets initiated when a person needs a breakthrough. Here, the person tries to reframe the situation by replacing an anchor, usually the weakest one, whereas the first pathway tries to build on the weakest anchor. For

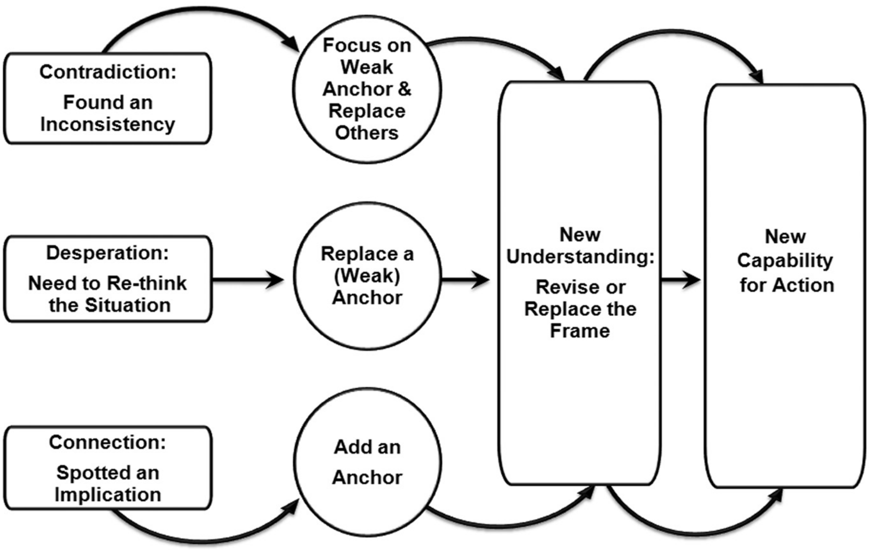

Figure 1. Anchor model of insight.

example, when Wagner Dodge realized that he could not outrun the wildfire in the Mann Gulch incident, his understanding of the situation was based on the steep uphill terrain, which he could not change, and the speed of the onrushing wall of flames, which he could not change. Another anchor for his frame was the combustible grass he was running though. And he could change that. He could set it alight, creating an escape fire that burned the fuel in front of him so that he could dive into its ashes and be saved. Most laboratory studies that set up impasses rely on this creative desperation pathway.

The third pathway gets initiated when a person sees a connection, often involving a new data element, that allows reframing, as when Darwin read Malthus's work on populations and realized that competition for resources could explain how species evolved. Here, the person adds a new anchor to an existing frame and then works out the implications. The few incidents in which the person was *not* surprised tended to follow this path. The person identified a leverage point and explored how that could open up a new strategy. The example cited earlier of Napoleon's tactics at Toulon was consistent with this pathway. The connection path does not necessarily involve a rejection of any of the initial beliefs, whereas the contradiction and breakthrough paths do depend on revising and/or rejecting initial beliefs.

Figure 1 reflects the finding that insights are not just about revising a person's understanding of situations. They can also be about how to act differently. In our sample, 54 of the 120 incidents included an insight about affordances that had not previously been identified.

Table 3 suggests that the contradiction path is fairly robust; 45 of the insights were triggered by contradictions. The creative desperation path seems less common; only 29 of the insights were generated in response to an impasse. The connection path was the most frequent of all, with 98 incidents, and overlaps the other two. The process of noticing contradictions or replacing weak anchors includes the forging of new connections as the frame is revised or a new frame is constructed.

This Anchor model of insight may have more utility than previous stage accounts of insight such as Wallas's (1926) four-stage model (preparation, incubation, flash of illumination, verification). In most of the incidents we examined, people did not deliberately prepare themselves prior to the insight and preparation was irrelevant for the insights that occurred by accident. An incubation stage was rarely called out and was implausible in many of the cases. The flash of illumination stage seems underspecified as well as incompatible with our finding of gradual insights.

We hope that the naturalistic study of insight will be useful for researchers who want to investigate the phenomenon further. We also hope that our findings will help organizations in their efforts to increase insights and innovations. One speculation is that there may be different approaches to support the pathways shown in Figure 1. The contradiction pathway may be supported by encouraging mental simulations of the implications of weak anchors rather than closing off such flights of fantasy to explain away inconvenient data. The desperation pathway may be supported by reviewing assumptions to reverse any that might be fostering fixation. The connection pathway might be supported by arranging for a flow of different concepts to create some turbulence.

Regardless of the approach pursued, we advise organizations against trying to establish a set of procedures to promote insights. The desire for procedures seems to be incompatible with the accidental quality of insights and reinforces the organizational goals of standardizing performance and reducing errors. We would encourage organizations to review whether their efforts to reduce errors may be interfering with insights and innovations.

Acknowledgments

The authors would like to thank Robert Hoffman for his helpful comments on several drafts of this article.

References

Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *Canadian Journal of Statistics*, *27*, 3–23.

Berkun, S. (2010). *The myths of innovation.* Sebastopol, CA: O'Reilly Media.

Berlinger, N. T. (1996). Vital signs: The breath of life. *Discover*, *17*(3), 102–104.

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences*, *9*(7), 322–328.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*, 1–49.

Cohan, W. (2009). *House of cards: A tale of hubris and wretched excess on Wall Street.* New York, NY: Doubleday.

Colligan, D. (2009, November). The Discover interview: Jocelyn Bell Burnell. *Discover*, pp. 66–74.

Crouch, T. (1989). *The Bishop's boys: A life of Wilbur and Orville Wright*. New York, NY: Norton.

Doyle, J. K., & Ford, D. N. (1998). Mental models concepts for system dynamics research. *System Dynamics Review*, *14*, 3–29.

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

Feltovich, P. J., Coulson, R. L., & Spiro, R. J. (2001). Learners' (mis)understanding of important and difficult concepts: A challenge to smart machines in education. In K. D. Forbus & P. J. Feltovich (Eds.), *Smart machines in education* (pp. 349–376). Menlo Park, CA: AAAI/MIT Press.

Gilhooly, K. J., & Fioratou, E. (2009). Executive functions in insight versus non-insight problem solving: An individual differences approach. *Thinking and Reasoning*, *15*, 355–376.

Hargadon, A. (2003). *How breakthroughs happen: The surprising truth about how companies innovate.* Cambridge, MA: Harvard Business School Press.

Heuer, R. J., Jr., & Pherson, R. H. (2010). *Structured analytic techniques for intelligence analysis.* Washington, DC: CQ Press.

Ippolito, M. F., & Tweney, R. D. (1995). The inception of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 433–462). Cambridge, MA: MIT Press.

Johnson, S. (2006). *The ghost map: The story of London's most terrifying epidemic—and how it changed science, cities and the modern world*. New York, NY: Riverhead Books.

Johnson, S. (2010). *Where good ideas come from: The natural history of innovation*. New York, NY: Riverhead Books.

Judson, H. F. (1996). *The eighth day of creation: Makers of the revolution in biology*. Cold Spring Harbor, NY: Cold Spring Harbor Press.

Klein, G. (2005). *The power of intuition*. New York, NY: Doubleday.

Klein, G. (2011). Critical thoughts about critical thinking. *Theoretical Issues in Ergonomics Science, 12*, 201–224.

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*, 462–472.

Klein, G., & Hoffman, R. (2008). Macrocognition, mental models, and cognitive task analysis methodology. In J. M. Schraagen, L. G. Militello, T. Ormerod, & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 57–80). Hampshire, UK: Ashgate.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, *21*(5), 88–92.

Klein, G., Phillips, J. K., Rall, E., & Peluso, D. A. (2007). A data/frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context* (pp. 113–155). Mahwah, NJ: Lawrence Erlbaum.

Liedtka, J., Rosen, R., & Wiltbank, R. (2009). *The catalyst: How you can become an extraordinary growth leader.* New York, NY: Crown Business.

Luchins, A. S., & Luchins, E. H. (1959). Rigidity of behavior: A variational approach to the effect of Einstellung. Eugene, OR: University of Oregon Books.

Maclean, N. (1992). *Young men and fire*. Chicago, IL: University of Chicago Press.
Man, J. (2002). *Gutenberg: How one man remade the world with words*. New York, NY: John Wiley.
Markopolos, H. (2010). *No one would listen: A true financial thriller*. Hoboken, NJ: John Wiley.
McGrath, R. G., & MacMillan, I. C. (2009). *Discovery-driven growth: A breakthrough process to reduce risk and seize opportunity*. Boston, MA: Harvard Business School Press.
Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*, 238–246.
Ogle, R. (2007). *Smart world: Breakthrough creativity and the new science of ideas*. Boston, MA: Harvard Business School Press.
Perkins, D. (2000). *Archimedes' bathtub: The art and logic of breakthrough thinking*. New York, NY: Norton.
Shilts, R. (1987). *And the band played on: Politics, people, and the AIDS epidemic*. New York, NY: St. Martin's.
Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.
Smith, S. M. (1995). Getting into and out of mental ruts: A theory of fixation, incubation, and insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 229–251). Cambridge, MA: MIT Press.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: implications for the rationality debate. *Behavioral and Brain Sciences*, *23*, 645–726.
Sternberg, R. J., & Davidson, J. E. (Eds.). (1995). *The nature of insight*. Cambridge, MA: MIT Press.
Topolinski, S., & Reber, R. (2010). Gaining insight into the "aha" experience. *Current Directions in Psychological Science*, *19*, 402–405.
Wallas, G. (1926). *The art of thought*. New York, NY: Harcourt Brace.
Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157–196). Cambridge, MA: MIT Press.
Wertheimer, M. (1945). *Productive thinking*. New York, NY: Harper.

Gary Klein is a senior scientist at MacroCognition LLC. He received his PhD from the University of Pittsburgh in 1969. He taught at Oakland University in Rochester, Michigan, from 1970 to 1974, before becoming a research psychologist at the Air Force Human Resources Laboratory at Wright-Patterson AFB in 1974. He started Klein Associates in 1978; it was acquired by Applied Research Associates in 2005.

Andrea Jarosz is a research assistant at MacroCognition LLC. She received her BS in psychology from Wright State University in 2010.