Defining Insight for Visual Analytics

Remco Chang, Caroline Ziemkiewicz, Tera Marie Green, and William Ribarsky, *University of North Carolina at Charlotte*

Many have argued that providing insight is the main goal of information visualization. Stuart Card, Jock Mackinlay, and Ben Shneiderman declare that "the purpose of visualization is insight,"[1] while Jim Thomas and Kris Cook propose in *Illuminating the Path* that the purpose of visual analytics is to enable and discover insight.[2] The idea that visualization should lead to insight seems logical, but researchers in the community have been slow to build on the concept because insight is difficult to define.[3-6] As Ji Soo Yi and his colleagues point out, although a few definitions of insight exist, no commonly accepted definition has emerged in the community.[6]

Interestingly, the visualization community is not the only one investigating insight. For the past two decades, researchers in cognitive neuroscience have been studying their own version of insight by examining neural activity. In their discipline, insight is a less ambiguous term. It specifically refers to what is commonly called an "aha" or "eureka" moment.[7] In fact, it is now possible to observe and identify when his or her is having such a moment by examining his or her neural activity.

It is clear that the scope of definitions of insight in the visualization community differs from that of the cognitive community. It appears that the visualization definitions of insight are generally broader but more vague than those in cognitive science. For example, Chris North categorizes insight to be "complex, deep, qualitative, unexpected, and relevant,"[3] which overlaps with the neurological definition. However, North and his colleagues also define insight as "an individual observation about the data by the participant, a unit of discovery,"[5] which does not bear any clear relation to the strict aha moment of cognitive science. Instead, it implies a focus on knowledge-building not found in the cognitive definition.

We suggest that what the visualization community defines as insight actually has two parallel meanings: a term equivalent to the cognitive science definition of insight as a moment of enlightenment, and a broader term to mean an advance in knowledge or a piece of information. We argue that for information visualization and visual analytics to provide and enable insight, both definitions need to be considered. But we must clarify and distinguish these definitions in order to develop methods to measure insight and evaluate visualizations.

Insight in Cognitive Science

The cognitive science community has used the term insight "to name the process by which a problem solver suddenly moves from a state of not knowing how to solve a problem to a state of knowing how to solve it."[8] (To distinguish this insight from the type of insight visualization researchers describe, we will call it *spontaneous insight*). In this tradition, spontaneous insight is a type of problem solving and differs from normal problem solving in several key ways. First, spontaneous insight doesn't appear to be facilitated by gradual learning heuristics such as bottom-up inductive reasoning. In fact, researchers have observed that focused effort on normal problem solving often inhibits spontaneous insight. Spontaneous insight usually occurs when a person is in a relaxed state[9] (such as when taking a shower in the morning). Second, whereas gradual problem solving requires no special inducement other than presenting someone with a problem, what precipitates spontaneous insight is still being discussed. One commonly held theory is that spontaneous insight often occurs when a person tries to solve the problem in a habitual way, fails, momentarily becomes frustrated (perhaps owing to incorrect assumptions or some other cognitive fixedness), mentally reorganizes the pieces of the puzzle (perhaps by breaking through a failed thought paradigm), and "suddenly" sees the solution.[8] Finally, in normal problem solving the path taken to the solution is conscious and logically clear to the problem solver; however, participants who experience a spontaneous insight often can't describe the thought process that led to it,[10] indicating that this insight occurs subconsciously and isn't a process that can be directly controlled, manipulated, or repeated.

In real-world situations, predicting when spontaneous insight will be used instead of traditional problem-solving is difficult. So, much spontaneous-

boot   boot camp
summer   summer camp
ground   campground

(a)   **(b)**

insight research in cognitive science has utilized carefully crafted laboratory problems that test subjects will most likely solve with sudden breakthroughs. One such problem is the nine-dot problem, which asks the participant to use four lines to connect nine dots that are arranged in a 3 × 3 matrix, without pulling the pen off the paper (see Figure 1a). Participants often fail at first by assuming their lines can't extend outside the matrix boundaries.[8] They seem to succeed when they relax their paradigms and try a different heuristic.

Similarly, Edward Bowden and his colleagues have developed a set of *compound remote associate* (CRA) problems for studies of spontaneous insight.[10] These present a list of three words to a participant, who then must come up with another word that makes a compound phrase when placed at either the beginning or end of each stimulus word. For example, the problem words "boot," "summer," and "ground" lead to the solution "camp" (see Figure 1b). Such problems can be solved through normal trial-and-error problem solving, by systematically coming up with potential solutions that work for one problem word and testing them against the other two. However, about 50 percent of the time, participants report that they identify the solution in a flash after several failed attempts.[9]

With the help of electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI), researchers can now observe participants' neural activities during experiments such as the nine-dot or CRA problems.[10] Two distinct patterns are observed that correspond to normal and insight-based problem solving. In both cases, frontal lobes, which are associated with working memory and executive processes, and temporal lobes, which hold long-term memory and semantic information, both show high activity.

However, in normal problem solving, the activity in the temporal lobe is continuous and localized mostly in the left hemisphere, which is thought to encode more detailed information in tightly related semantic networks. This indicates that normal problem solving involves a narrow but continuous focus on information highly relevant to the problem at hand. In contrast, when participants solve a problem with spontaneous insight, the right temporal lobe shows a sharp burst of activity, specifically the superior temporal gyrus.[10] Unlike the left temporal lobe, the right temporal lobe is thought to encode information in coarse, loosely associated semantic networks. This suggests that spontaneous insight occurs through sudden activation of less clearly relevant information through weak semantic networks, which corresponds to a participant's paradigm shift following an impasse.
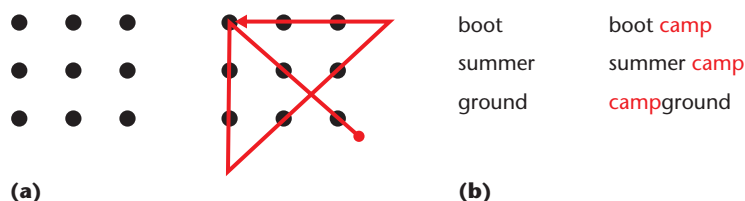
These findings suggest that spontaneous insight is qualitatively different from everyday problem solving. It involves a unique pattern of neural activity that corresponds with the unique sensation of the "aha" moment that participants report. However, although cognitive scientists have successfully identified the neural patterns of the spontaneous-insight phenomenon and can now observe and measure the insight process, they've had little success in understanding why spontaneous insight occurs in normal situations. More important, aside from laboratory experiments using artificially designed simple problems, cognitive scientists don't understand how to promote spontaneous insight for solving complex real-world problems like the ones the visualization community encounters regularly.
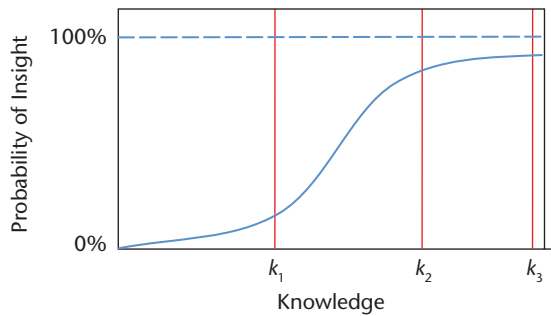
Insight as Knowledge and Information

Recently, Yi and his colleagues provided a comprehensive survey on information visualization literature that considered insight as a goal or a measurement.[6] On the basis of sensemaking theories, they concluded that four distinct but intertwined processes in visualization can lead to insight: provide overview, adjust, detect patterns, and match mental model. Although these processes can in some cases result in an aha moment, it's clear that their research as well as the literature used in their survey considers insight more or less as units of knowledge. (For this discussion, we call this type of knowledge-building or model-confirming insight, simply, "insight.")

For example, one surveyed paper is a case study by Nina Amenta and Jeff Klingner on visualizing evolutionary trees. They state that "our tool allows the biologists to interactively visualize and explore the whole set of trees, providing insight into the overall distribution and possible conflicting hypothesis."[11] In this context, insight refers to knowledge about the overall distribution. Switching the word "insight" with "knowledge" doesn't effectively alter the statement's meaning.

Similarly, Victor Gonzales and Alfred Kobsa[12] report the result of their study on adopting an information visualization system by administrative data analysts: "the analyst determined the answers to these questions, but also came up with further insights that she shared with people from other administrative units. She used the discovered information to advise other administrators of certain

Figure 2.
The possible
relationship
between
knowledge and
insight. We
hypothesize
that a
nonlinear
relationship
exists between
the amount of
knowledge a
user gains from
using visual
analytics and
the probability
that the user
might have a
spontaneous
insight on the
task at hand.

previously unknown relationships in their data." The two sentences mirror each other and indicate that the insight shared by the analyst is in fact information of previously unknown relationships (or knowledge) in the data that she discovered.

These two examples are telling. Insight in information visualization and visual analytics not only has a different sense than in cognitive science but also is considered entirely differently. In the visualization community, researchers often talk about discovering insight, gaining insight, and providing insight. This implies that insight is a kind of substance, and is similar to the way knowledge and information are discussed. In the cognitive science community, researchers more often discuss experiencing insight, having an insight, or a moment of insight. In this context, insight is an event. The fact that the two definitions refer to entirely different kinds of concept is a major impetus to consider them separately.

Is the Goal of Visualization Insight?

With a clearer definition of insight, the statement that the goal of visualization is insight must be re-examined. On the basis of the cognitive definition of insight, this statement restricts visualization into considering only a specific mode of problem solving that produces results that, although measurable, aren't easy to track. On the other hand, considering insight only as knowledge or information limits visualization's potential to structured knowledge building and information display.

Intuitively, visualization ought to promote both types of insight. However, the differences between the two make this goal seemingly difficult. Spontaneous insight is a form of problem solving used to solve difficult, seemingly incomprehensible problems. Knowledge-building insight, on the other hand, is a form of learning that builds a relationally semantic knowledge base through a variety of problem-solving and reasoning heuristics. Additionally, spontaneous insight seems to depend on the desertion of applied paradigms and schematic structures, whereas knowledge-building insight generally depends on schematic structures (such as a mental model) to find patterns as well as to infer.

Although the goals of promoting spontaneous insight and knowledge-building insight appear disparate, we propose that in fact they're related. As little as is known about the origin of spontaneous insight, it doesn't arise out of thin air. If spontaneous insight comes from the unexpected reconfiguration of semantic knowledge,[10] then relevant knowledge about a problem must be necessary for spontaneous insight to arise. This is evident in the nine-dot and CRA experiments, during which participants are often observed to experience spontaneous insight only after attempting to solve the problem using normal problem-solving methods. We also can see that this is true more generally and for deeper insights: Einstein didn't come up with the theory of relativity out of thin air but rather based it on experiments inconsistent with existing theories and previous mathematical work. Conversely, the major paradigm shifts associated with spontaneous insight can create new structures and relationships in a user's understanding of a problem, which can then serve as the schematic structures needed for generating future knowledge-building insights.

We propose that a similar relationship between knowledge building and spontaneous insight can be found in using visualization to solve complex problems. The existence of deep, complex knowledge about a subject increases the likelihood someone can make a novel connection within that knowledge. Likewise, each major spontaneous insight opens up the possibility of new directions for knowledge building. Together, the two types of insight support each other in a loop that allows human learning to be both flexible and scalable. Because we know more about designing systems for knowledge building than for the more elusive goal of promoting spontaneous insight, we can focus on the former and rely upon this proposed relationship to ultimately encourage both kinds of insight.

Figure 2 shows our hypothesis of how using an exploratory visual analytical tool to increase the user's knowledge could increase the probability of spontaneous insight. We posit a general positive correlation between the two, but the relationship isn't linear. As Figure 2 shows, when the user has only a limited amount of knowledge (0 to k_1), spontaneous insight won't likely occur. As the amount of knowledge increases (k_1 to k_2), the probability of spontaneous insight increases sharply. Finally, after a certain point (k_2 to k_3), further increase of knowledge increases the probability in only a limited fashion until it's asymptotically close to a spontaneous insight occurring. On the other hand, a reduction in the probability of gaining a spontaneous insight undoubtedly occurs, at least for a while, if the user is distracted from this freer knowledge association.

Although this model is simple, we believe that it can provide a starting framework for more accurate insight-based evaluation of visualizations. But whatever model is chosen, our main point is that spontaneous and knowledge-building insights should be considered distinct because the best approaches to gain one or the other are different. For spontaneous insight, we can evaluate exploratory, "prequery" approaches that keep one "in the cognitive zone" or "in the flow," and quantitatively identify when a spontaneous insight occurs through an EEG or fMRI. For knowledge-building insight, we can evaluate detailed knowledge-gathering methods and look to appropriate user studies to measure how much knowledge a user gains. Using these combined approaches, we can not only more accurately determine visualization tool's effectiveness, but also provide cognitive scientists with more complex problem-solving artifacts (they have few available) and shed light onto how to promote the two types of insight through visualization tools to solve real-world problems.

References

1. S.K. Card, J.D. Mackinlay, and B. Shneiderman, eds., *Readings in Information Visualization: Using Vision to Think*, Morgan Kaufmann, 1999.
2. J. Thomas and K. Cook, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, Nat'l Visualization and Analytics Center, 2005.
3. C. North, "Toward Measuring Visualization Insight," *IEEE Computer Graphics and Applications*, vol. 26, no. 3, 2006, pp. 6–9.
4. C. Plaisant, J.-D. Fekete, and G. Grinstein, "Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository," *IEEE Trans. Visualization and Computer Graphics*, vol. 14, no. 1, 2008, pp. 120–134.
5. P. Saraiya, C. North, and K. Duca. "An Insight-Based Methodology for Evaluating Bioinformatics Visualizations," *IEEE Trans. Visualization and Computer Graphics*, vol. 11, no. 4, 2005, pp. 443–456.
6. J.S. Yi et al., "Understanding and Characterizing Insights: How Do People Gain Insights Using Information Visualization?" *Proc. 2008 Conf. beyond Time and Errors* (BELIV 08), ACM Press, 2008, pp. 1–6.
7. J. Lehrer, "The Eureka Hunt: Why Do Good Ideas Come to Us When They Do?" *The New Yorker*, 28 July 2008, pp. 40–45.
8. X. Mai et al., "'Aha!' Effects in a Guessing Riddle Task: An Event-Related Potential Study," *Human Brain Mapping*, vol. 22, no. 4, 2004, pp. 261–270.
9. J. Kounios et al., "The Prepared Mind: Neural Activity Prior to Problem Presentation Predicts Subsequent Solution by Sudden Insight," *Psychological Science*, vol. 17, no. 10, 2006, pp. 882–890.
10. E.M. Bowden et al., "New Approaches to Demystifying Insight," *Trends in Cognitive Sciences*, vol. 9, no. 7, 2005, pp. 322–328.
11. N. Amenta and J. Klingner, "Case Study: Visualizing Sets of Evolutionary Trees," *Proc. 2002 IEEE Symp. Information Visualization* (INFOVIS 02), IEEE Press, 2002, pp. 71–74.
12. V. González and A. Kobsa, "Benefits of Information Visualization Systems for Administrative Data Analysts," *Proc. 7th Int'l Conf. Information Visualization* (IV 03), IEEE CS Press, 2003, pp. 16–18.

Remco Chang *is a research associate in the Department of Computer Science at the University of North Carolina at Charlotte and a member of the Charlotte Visualization Center. Contact him at rchang@uncc.edu.*

Caroline Ziemkiewicz *is a PhD candidate in computer science at the University of North Carolina at Charlotte. Contact her at caziemki@uncc.edu.*

Tera Marie Green *is a PhD student and research assistant at the School of Interactive Arts and Technology, Simon Fraser University. Contact her at grepmon@gmail.com*

William Ribarsky *is the Bank of America Endowed Chair in Information Technology at the University of North Carolina at Charlotte and the founding director of the Charlotte Visualization Center. Contact him at ribarsky@uncc.edu.*

Contact editor Theresa-Marie Rhyne at tmrhyne@nscu.edu.