The Current State of Insight Research

Leo Trottier

University of Toronto

299-73 St. George Street

Toronto, ON M5S 2E5

Abstract

Examines existing definitions of insight and insight problems, revealing a lack of consensus and clarity, subsequently proposing what would be necessary in a legitimate definition of an insight problem and possible means of reaching one. The lack of consensus implies that experimental uses of insight problems lack validity, since the current *de facto* determiner of what constitutes an insight problem is not the problem itself but the effect it produces. So-called 'insight' problems are seen to be an inconsistent set where multiple undefined influences may be at play in causing the requisite 'Aha!' experience. It becomes apparent that a detailed analysis of 1) what is insight, 2) what constitutes an insight problem, and 3) how we can create new insight problems, is necessary for the field to move forward.

The Current State of Insight Research

The phenomenon of insight was first proposed by Gestalt psychologists around the first quarter of the 20th century. In the Gestalt theory of mind, insight was described as the sudden finding of a solution to a particular problem, which can be attributed to successful problem restructuring. However, without a functional definition of this restructuring, it provides neither a useful definition of the phenomenon, nor does it facilitate an explanation for how it occurs. Breaking insight down into its definition and explanation is important and can be justified on the grounds that it helps to avoid circularity (Schooler, 1995). Without the distinction, insight remains in a constant state of flux, as its meaning is closely dependent on its mechanism. In fact, some psychologists still consider the term ‘insight’ to be entirely meaningless (Dominowski & Dallob, 1995). Attempts to describe and explain this phenomenon since the Gestalt period have nevertheless continued. This research, however, has been plagued by the original faulty assumptions of Gestalt psychologists, and their use of dated and arbitrarily chosen ‘insight problems’. The crucial obstacle confronting insight researchers could thus be seen as the inadequacy of the definition of insight and insight problems. This has caused the field of insight research to grind to a halt. To proceed any further, a new perspective is needed on what true insight problems are as well as a theoretical foundation for their creation. Without these, there is little hope of reviving this ailing field of cognitive science.

An adequate definition of insight should allow for the invention of insight problems in an exact and predictable manner. It should be a set of necessary and sufficient conditions for a problem to belong to the class of insight problems. A rigorous classification of insight problems has been stressed, because at present problems are normally chosen from previous studies without regard for what makes the problem truly an insight problem (Wieth & Burns, 2000).

Worse yet, it seems as though little thought has been given as to how any two insight problems involve similar processes.

One useful view of insight gleaned from the current literature is a simple one, but a look at its development is needed to appreciate its functionality. Richard Mayer's definition, that it is the '[sudden move]' from a state of not knowing how to solve a problem to a state of knowing how to solve it" (1995), is a good place to start; however, the consideration of other current definitions leaves the reader, in most cases, perplexed by their vagueness. Mayer's abstraction of the Gestalt view of insight describes insight as 1) completing a schema, 2) the reorganisation of visual information, 3) the reformulation of a problem, 4) the overcoming of a mental block, and/or 5) the finding of a problem analogue. Thus the question becomes what problems, if any, are *not* insight problems? In truth, these definitions of insight are relatively useless; there is no way that the above views are definitions in the more rigorous sense.

The first view, completing a schema, is described as the filling in of a gap in a coherent set of data. This does little to help explain insight, and fails to help classify insight problems, as this definition requires an understanding of what a schema is, and what it is to complete it (Mayer, 1995).

The second view, that insight is the reorganisation of visual information, does slightly better: in the specific case of visual problems, it is the redistribution of the problem's components. However, this involves the implicit assumption that the problem's components can be organised in the first place (a problem of relevance), and that they can be subsequently reorganised (a changing in the structure of the relevant parts). That is, as in many definitions, the terms used are in themselves imprecise (Mayer, 1995).

The third view, that of reformulation of a problem, was proposed by Gestalt psychologist

Karl Duncker. Essentially, it is the redefinition of the goal, or the redefinition of the given information. We can now begin applying this definition to proposed problems. That is, a problem is an insight problem if a solver's initial approach to it cannot solve it. Hence, a change in approach provides the method of obtaining the solution. This description, while more robust, nonetheless does not adequately involve the solver in its definition because it fails to consider the suddenness associated with an instance of insight.

This suddenness can be explained, in part, by the fourth perspective of insight as the removal of mental blocks – for example, impasses caused by functional fixedness. In the removal of a mental block we '[escape] the tyranny' (Gick & Lockhart, 1995) of our original interpretation, and overcome the impasse. We then experience exuberance because we have seen the problem's gimmick, and avoided it. Unfortunately, this definition of insight implies that solutions are already in the mind, waiting to be revealed, and that if the solver could remove the barrier they would achieve insight. Such is often not the case, as solvers frequently note that it was the lack of a key piece of knowledge, (Seifert *et al*, 1995) and not a block *per se* that prevented insight.

Finally, the notion of finding a problem analogue serves as perhaps the best of the five quasi-definitions of insight – namely, during problem solving we experience insight when we discover an analogous problem, or we see that the method of how to solve one problem can be applied to the current one. This view does the best job at both defining insight, and explaining the phenomenon of suddenness, that is, when an analogue is found, the method of solution appears quickly, because it was already there in the analogous one. All that need happen is a mapping of the problem elements. Nevertheless, this idea of finding a problem analogue is stretched when applied to many so-called insight problems. The finding of an analogous problem appears to be more appropriate as an explanation for insight (Chen & Daehler, 2000) rather than a definition. Analogous problems can aid in the formulation and interpretation of future problems, reducing the need for insight, and causing similar problems to become more routine (Chen and Daehler, 2000).

So perhaps a simpler, less contrived definition is what is needed. Seifert *et al* (1995) provide such a definition in quoting *Webster's New World Dictionary*, namely 'seeing and understanding the inner nature of things clearly, especially by intuition'. This description, however, is plagued by the most serious problem of the previous ones: it is far too general. It would be difficult to apply to the classification of problems, and is completely reliant on the concept of 'understanding', an equally obscure concept.

Alternatively, Steven Smith (1995) proposes to define insight by making a distinction between '*insight*, *insight experience*, and *insight problem*' (emphasis in original). This is a useful approach, for it may help to reduce 'insight' to its constituent elements. However, even with this more precise approach, he defines insight as 'an understanding'. He goes on to define the insight experience as 'the sudden emergence of an idea into conscious awareness, the "Aha!" experience' (Smith, 1995). Finally, in what one hopes would be a *coup de grace*, he defines the insight problem as 'one for which the solution is more likely to be reached via an insight experience'. One cannot help but feel let down.

These definitions do not satisfy the criteria of usefulness and precision. While insight is often described as an 'understanding', insight problems are likewise defined subjectively, in reference to an experience felt by subjects when finding the solution to a problem. Thus, the problems used to study insight are not problems that have been developed theoretically but instead by trial and error. Problems are found (mainly from previous studies, see Weisberg, 1995, pp 184-191), and subsequently classified as insightful by the Aha! effect they have when their solution is found, rather than the other way around. As Robert Weisberg notes, this method shows that ‘the classification system has no theoretical grounding’ (1995).

Other definitions of insight and classifications of insight problems are equally obfuscatory and circular (*e.g.* Sternberg and Ben-Zeev, 2001 or Gilhooly, 1988). This causes us to return to our first, and what will be called ‘simple’ definition of insight. It is ‘a transition event in 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”’ (Schooler *et al*., 1995). The word *how* is key to this definition. Perhaps, instead of insight problems being a set of problems that use some special process (for example Seifert *et al*., 1995), they are merely problems for which, following initial attempts to solve the problem, the solver realises they do not know *how* to solve it. One may even remove ‘sudden’ from the definition. In other words, insight is the discovery of a method to solve a particular problem where one was not previously known.

So how does this definition account for the suddenness or ‘Aha!’ experience? In problems in which insight does not occur, the solver either already knows how to solve the problem, or fails to solve it. However, in the group that tries to solve a problem persistently, the goal and solution states of the problem become very familiar. Thus, they develop a good understanding of their impasse: they do not know how to solve the problem, but are aware of the missing link. When they finally think of the alternate approach (how they do is an issue of explanation) they are already familiar with their goal, and so can easily solve the problem. This causes the solution to appear suddenly and assertively: the solver is certain that they now know how to solve the problem, thus leading to a feeling of contentment.

Hence, this emphasis on *how* to solve rather than *what* to do once *how* is known is what sets insight problems apart from conventional problems. In conventional or ‘routine’ problems, the solution method is already obvious, and the solver has merely to proceed through the steps. Purely insight problems might then be referred to as meta-problems, for they are problems on a higher level. For the purest insight problems, the steps are trivial (unlike, for instance, mathematical integration), and it is this ‘meta’ aspect of the problem that must be overcome. This is why problems such as the mutilated checkerboard problem (Appendix A, 1), are easily solved once a different method (other than brute-force confirmation) is applied.

In fact, this interpretation of insight can be applied to many classical insight problems. The nine-dot problem (Appendix A, 2) can now be seen to be a problem where the initial or intuitive method of solution is to stay within the ‘box’ created by the nine dots. This explains why, when insight is experienced, the solution is still not instantaneous, and steps must be worked out even after insight occurs. Thus, instructing subjects that the proper solution method was to leave the square matrix ‘did not significantly facilitate solution’ (Weisberg, 1995). This can now be attributed to the subject’s having understood the method, but still needing to iterate through the steps to arrive at the solution.

Fixation can now be attributed to an inability to form a new method of solution, rather than an impasse in a correct solution method. Thus, functional fixedness on the typical use of pliers in the classical two-string problem (Appendix A, 7) causes an inability to see that the correct solution method is to cause the strings to move toward the subject, rather than the other way around. The steps involved once the correct method is established are now trivial.

This definition also fits nicely with the interpretation that insight is ‘restructuring’. Seeing a problem in a new way, is, at its essence, looking at a problem and seeing a different method for solving it. The concept of the reformulation of a problem is now taken care of, as it is simply a change in how a solver approaches a problem, i.e. a different method of solution.

Another benefit of the simple definition is that it gives a deeper perspective on the empirical warmth ratings performed by Metcalfe (Davidson, 1995). Whereas Metcalfe had subjects judge their warmth on a scale of zero to ten for both insight and incremental problems, according to the simple definition of insight, this would lead to a mismatch between the two graphs (Figure 1). Thus, in incremental problems, the subjects would be starting their solving at a different, more advantageous state than in insight problems: they already know how to solve it. If the warmth rating had a value representing when the subject was confident in their method (five, say), it would be seen that in incremental problems, the subject has a higher rating from the get-go. This would be seen in contrast to insight problems, where the subject's progress in developing a method of solving the problem would be charted below the value of five.

The insight problem and incremental problems presented above are the types of problems given in many studies of insight problems. The incremental problem is typically a routine problem that the subject can confidently solve, but involves a number of steps. The insight problem, on the other hand, is one in which once a method for finding the solution is found, the steps needed to solve the problem are few (thus the quickness once the method was found). Interestingly, a method for classifying insight problems might be a test of its conduciveness to having the answers given in a multiple-choice fashion. By the above criteria, pure insight problems would involve no 'working out', and thus an answer, once shown, would appear self-evident. Insight problems may have the important property of being unable to test by multiple-choice.

This analysis in its own right suggests a different method of analysing potential problems, namely: the problem should be looked at in terms of the difficulty in developing the solution method, as well as another examination of the number of steps needed to solve the problem once it has been correctly formulated. This implies a radical concept: perhaps problem difficulty can be represented in two dimensions, where one axis represents the steps involved in solving a problem, and the other represents the meta-problem aspect, or finding the way in which to solve the problem. The routine method dimension involves effable problem solving, while the other dimension involves ineffable problem solving (Wieth & Burns, 2000). Although the experimental implications of this approach are tantalising, further speculation would be aimless.

Applying the dimensional perspective to problems such as the nine-dot problem is congruent with the findings of a number of researchers: the problems involve two aspects, and thus could be called hybrid problems (Weisberg, 1995). On the other hand, problems such as the Dunker radiation problem (Mayer, 1995) could be seen to be nearly purely insight, as the number of steps involved once multiple converging rays have been considered are very few. Furthermore, other problems that are normally seen as insight problems become, instead, exploits of interpretation (Weisberg, 1995), with little connection to reality. As Weisberg puts it, to say that 'one has "solved" the animals in pens (Appendix A, 3) problem by building four concentric pens, with all the animals in the middle one, violates the meaning of the phrase *four pens*.' He calls the solution, instead, one pen with four fences. The problem ceases to be an insight problem, and becomes instead a play on words, with all the difficulty now attributable to rationalisation and common diction. The solution involves a distorted interpretation that is not found in normal usage (Weisberg, 1995).

The same applies to the checker games problem (Appendix A, 4), as two men would never be referred to in such a way in natural speech without the implication that they were playing against each other. Similarly for the lazy policeman problem (Appendix A, 6). Driving was implied by the extra information, since references to one-way streets and rail crossings would not be made for a pedestrian in the first place. For the Charlie problem (Appendix A, 5), the issue is more one of a lack of information, as in reality Charlie's cause of death would have been plainly obvious (not to mention the fact that there is really no 'right' answer). This runs contrary to Gick and Lockhart's (1995) position that these artificial misinterpretations constitute pure insight. It is because these misinterpretations are contrived through deliberate misleading by the problem poser that the problems approach perversity and lose their ecological validity.

So how does one go about creating new insight problems? It seems that in actuality there are no problems that will require insight ability in everyone – for if the solver's original formulation of the problem is correct, the problem be solved in a routine fashion (Weisberg, 1995). As well, since it is possible to train subjects to perform better on certain types of insight problems (Kershaw & Ohlsson, 2001), the ability to generate new methods of problem solving will cause the classical problems to become useless. It thus seems unlikely that there is any such thing as a universal insight problem, since, for instance, while one person may have no idea how to solve the mutilated checkerboard problem, a mathematician may see it as a routine topology exercise.

The languishing state of insight research will not improve without a precise definition of insight, and more importantly, a robust and dynamic way of developing insight problems. Current research is inherently flawed because it uses problems that have little theoretical foundation. A new approach must be taken in all avenues of insight research, from empirical warmth-ratings to taxonomic systems for judging a problem's difficulty, contingent on a solver expertise. Without an exact definition of insight and insight problems, it will likely never be a rigorous sub-domain of cognitive science.

References

Chen, Z., & Daehler, M.W. (2000). External and Internal Instantiation of Abstract Information Facilitates Transfer in Insight Problem Solving. *Contemporary Educational Psychology, 25.* 423 – 449.

Davidson, J.E. (1995). The Suddenness of Insight. In R.J. Sternberg & J.E. Davidson (Eds.), *Nature of Insight.* (pp 125 – 155). Cambridge: MIT Press.

Dominowski, R.L. & Dallob, P. (1995). Insight and Problem Solving. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 33 – 62). Cambridge: MIT Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 68(5),* whole no. 270.

Finke, R.A. (1995). Creative Insight and Preinventive Forms. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 255 – 280). Cambridge: MIT Press.

Gick, M.L., and Lockhart, R.S. (1995). Cognitive and Affective Components of Insight. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 197 – 228). Cambridge: MIT Press.

Gilhooly, K.J. (1988). *Thinking: Directed, undirected, and creative.* San Diego: Academic Press

Kershaw, T.C. & Ohlsson, S. (2001). *Training for Insight: The Case of the Nine-Dot Problem.* Paper presented at the Twenty Third Annual Meeting of the Cognitive Science Society.

Mayer, R.E. (1995). The Search for Insight: Grappling with Gestalt Psychology's Unanswered Questions. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 3 – 32). Cambridge: MIT Press.

Perkeins, D.N. (1995). Insight in Minds and Genes. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 495 – 533). Cambridge: MIT Press.

Schooler, J.W., & Melcher, J. (1992) The Ineffability of Insight. In R.A Fink, T.B. Ward, & S.M. Smith, *Creative Cognition: Theory, research and application.* (pp 97-134). Cambridge: MIT Press.

Schooler, J.W., Fallshore, M., & Fiore, S.M. (1995). Epilogue: Putting Insight into Perspective. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 559 – 587). Cambridge: MIT Press.

Seifert *et al.* (1995). Demystification of Cognitive Insight: Opportunistic Assimilation and the Prepared-Mind Perspective. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight.* (pp 65 – 124). Cambridge: MIT Press.

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: MIT Press.

Sternberg, R.J. & Ben-Zeev, T. (2001) *Complex Cognition: The Psychology of Human Thought.* New York: Oxford University Press.

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: MIT Press.

Wieth, M. & Burns, B.D. (2000). *Motivation in Insight versus Incremental Problem Solving.*

Paper presented at the Twenty Second Annual Meeting of the Cognitive Science Society,

Philadelphia, Pennsylvania, August.

Appendices

Sample problems (Source: Weisberg, 1995, pp 184 – 191, except where noted)

1. **Mutilated Checkerboard problem:** An 8 x 8 chess board has two of its diagonally opposite corners removed. Would it be possible to fit dominoes on every space of the board, where a domino covers exactly two adjacent chess board squares? Why or why not?
   Solution: Once it is seen that every domino must cover a black and a white piece, the fact that two pieces of the same colour are removed implies that 31 dominoes must cover a field with 30 squares of one colour, and 32 of the other. But the dominoes always touch an even number of either colour, therefore the dominoes could not cover the whole board. (Gick and Lockhart, 1995)

2. **Nine-dot problem:** Without lifting your pencil from the paper, connect the nine dots by drawing four straight lines (Figure 2).

3. **Animals in pens:** Describe how you can put 27 animals into four pens so that there is an odd number of animals in each pen.
   Solution: Make four concentric pens and put all 27 in the center pen.

4. **Checker games:** Two men play five checker games and each wins an even number of games, with no ties. How is that possible?
   Solution: The men did not play against each other.

5. **Charlie:** Dan comes home from work and finds Charlie lying dead on the floor. Also on the floor are some broken glass and some water. Tom is in the room too. Dan takes on look around and immediately knows how Charlie died. How did Charlie die?
   Solution: Charlie, Dan's pet fish, died of lack of oxygen when Tom, Dan's cat, knocked over the fishbowl, causing it to shatter and spill its contents.

6. **Lazy policeman:** A woman did not have her driver's license, with her. She failed to stop at a railroad crossing, then ignored a one-way traffic sign and traveled three blocks in the wrong direction down the one-way street. All this was observed by a policeman, yet he made no effort to arrest the woman. Why?
   Solution: The woman was walking.

7. **Two String Problem:** Subjects must tie together the ends of two strings suspended vertically from a ceiling, even though the strings are widely separated and cannot be grasped simultaneously at the outset. A number of objects are provided, including a pair of pliers (typically).
   Solution: Often, subjects are fixated on the typical use of the objects, which is useless. However, the simplest and unintuitive solution is to attach an object (often the pliers) to one string to act as a pendulum. The subject swings the weighted string, and then takes the end of the non-weighted string to the middle of the space between the two strings, pausing to wait until the "pendulum" arcs back toward the subject (Seifert *et al*, 1995).

Figure Captions

Figure 1.

Figure 2. The nine-dot problem and its solution.

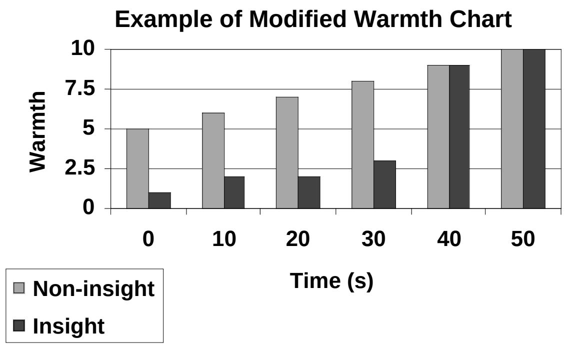

Figure 1.

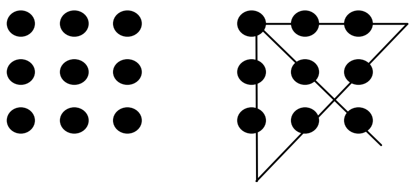

Figure 2.