Analogies Out of the Blue: When History Seems to Retell Itself

Lelyn Saner (lsaner@gmu.edu)
Department of Psychology, Applied Cognition Program
MSN 3F5; George Mason University
Fairfax, VA 22030-4444

Christian Schunn (cschunn@gmu.edu)
Department of Psychology, Applied Cognition Program
MSN 3F5; George Mason University
Fairfax, VA 22030-4444

Abstract

To explain the origins of new scientific ideas, historians and philosophers of science point to examples where scientists appear to have drawn analogies between their scientific domain and some very different domain. By contrast, research from the psychology lab suggests that those kinds of analogies are very difficult to obtain in even the simplest situations. To resolve this potential conflict, we examine the analogies that occur in psychology lab group and formal colloquium settings. This approach can be viewed as a cross-sectional approximation of an historical analysis. We find that as the setting moves further away from the original discovery, the way different types of analogies appear to be used changes. In particular, analogies between very different domains are never used in reasoning in the lab group, whereas they are frequently used in reasoning in formal colloquium presentations. Yet, we find that analogy between very similar domains remains an important source of new ideas and a method for solving problems in scientific settings.

Introduction

The research reported here has been done in response to a long-standing question in philosophy and psychology of science. From where do scientists get their ideas? Previous research on this question has suggested that analogy can be an important source of scientific ideas (e.g., Gentner, Brem, Ferguson, Markman, Levidow, Wolff, & Forbus, 1997; Gentner & Jeziorski, 1989; Thagard, 1988). For example, Kekulé is said to have developed the idea of a circular molecular structure for benzene from a dream of a snake swallowing its tail.

In addition, there is already much research on how analogies are used by people in various domains to explain, explore, infer, and persuade others with respect to a particular idea or approach to understanding. Some of these domains include politics (Blanchette & Dunbar, 1997), math education (VanLehn, 1998), history education (Young & Leinhardt, in press), narrative interpretation (Holyoak, 1985), and molecular biology (Dunbar, 1995).

A paradox arises from this previous work on analogy. On the one hand, when philosophers and historians of science have discussed the role of analogy in science, they have referred primarily to what one might call long-distance analogies— analogies between two highly distinct domains (e.g., Boden, 1993; Holyoak & Thagard, 1995; Koestler, 1964; Nersessian, 1985; Rhodes, 1986; Ueda, 1997). For example, ideas about the structure of the atom were said to derive from ideas about the structure of the solar system (i.e., several smaller bodies orbiting one larger body at the center, with much space between). Here, one is connecting a *source* in the domain of astronomy to a *target* in the domain of nuclear physics. Another example is the case of Kekulé and the analogy between benzene and a snake. Here the source domain is either mythology, everyday experience, or artistic imagery, and the target domain is chemistry. In each of these kinds of famous examples that are much discussed in history and philosophy of science, the source and target domains are almost completely unrelated (e.g., Boden, 1993; Koestler, 1964; Nersessian, 1985; Rhodes, 1986).

On the other hand, psychological research on analogy has found that even when cued, subjects will fail to connect two very analogous scenarios when they are from different domains. One of the most famous, and often cited, demonstrations of this result comes from Gick and Holyoak (1983). They showed that, after reading a story of how a general takes a large army across multiple bridges, subjects were still unable to see the application of it to an analogous medical problem. This result was initially quite surprising and has been frequently replicated: subjects usually do not spontaneously note long distance analogies (Holyoak & Koh, 1987; Mandler & Orlich, 1993; Ross, 1989).

Thus we have a paradox: If people do not spontaneously see long-distance analogies in such simple situations, how can scientists see long-distance analogies in the complexities of their research?

There are at least three simple resolutions to this problem. First, it is possible that scientists are more intelligent, creative, or insightful than the average psychology experiment subject, and thus are able to find and use long distance analogies. Certainly, individuals are thought to vary along many cognitive dimensions including creativity (Gardner, 1983).

Second, scientists have much deeper knowledge of their subject domain and this deeper knowledge base may be more supportive of long-distance analogies. Along those lines Novick (1988) found that students with greater math expertise were better able to make analogies involving math problems and other domains.

Third, it is possible that scientists do not make much use of long-distance analogies, and the historical record is misleading. Under this explanation, the psychological research is correct even for scientists: long-distance analogies are not typically used to solve problems. To account for long distance analogies described by historians of science, this explanation assumes that the historical record is misleading. In particular, it may be that the historical record contains descriptions of long distance analogies that were developed by the discoverers after the discovery was made. These analogies, while not used as the source of the discovery, may have been persuasive to the discoverer and were used to explain the discovery to others. In some cases, the scientists themselves have pointed to these long distance analogies as the source of their ideas. However, these accounts were typically long after the discovery, and psychological research has demonstrated that retrospective accounts of insights and analogies are frequently very inaccurate (Dunbar, 1996; Nisbett & Wilson, 1977; Schunn & Dunbar, 1996). This last possibility does not suggest that long-distance analogies are never used by scientists. Instead, it merely argues that long-distance analogies may not be as common or important as the previous accounts would suggest.

To resolve these issues, one can observe scientists as they are conducting their research, before the historical record can be rewritten. One such approach, developed by Kevin Dunbar (1995, 1996; Dunbar & Baker, 1994), is to analyze the discussions that occur in lab groups. Increasingly, science is being conducted in the context of a lab group, consisting of a head researcher, postdocs, graduate students, and research assistants. These lab groups meet on a regular basis and serve a very important function for the discovery process: new experiments are developed; data is interpreted; and alternative hypotheses are proposed.

Dunbar (1996) used this methodology to examine the kinds of analogies used by leading molecular biology lab groups as they made important discoveries. Analogies were coded along two dimensions: whether the analogy was within-organism, other-organism, or non-biological (or distant), and what the goal of the analogy was (form a hypothesis, design an experiment, fix an experiment, or explain a concept). Dunbar found that, of the 99 analogies identified, only two were from distant sources, and both were used to explain a concept rather than form a hypothesis or design or fix an experiment. What he concluded from his study was that distant analogies may not have a role in the making of discoveries, and that scientists use them more to explain a new concept to an audience. He also noted that scientists did not often remember analogies generated during meeting (including the within-organism analogies). This suggests that analogies may be used as scaffolding for the construction of new ideas, and are then discarded once they have served their purpose. This amnesia for the analogies that were used may explain the conflicting view of analogy from history of science. That is, the analogies that were actually used in making the discoveries are simply missing from the historical record, leaving only the analogies used in papers and presentations to convince others (i.e., the distant analogies).

To explore this resolution of the analogy paradox, we use a modification of Dunbar's methodology. We contrast the discussions that occur in lab groups with those that occur in colloquia. The assumption is that research is at an early stage when it is being addressed in lab groups. At this point, ideas are formulated and focused, preliminary studies are planned or data from them is analyzed to determine next steps. Since ideas are formulated and developed in this context, the kinds of analogies used here provides insight into whether long distance analogies are used as sources of ideas. In other words, will Dunbar's results generalize to other domains, in this case, that of psychology?

Colloquia, by contrast, present later-stage or completed studies. Conclusions and theories are typically well developed. Highlights from a sequence of studies rather than all the details of a single study are presented. Well-formulated ideas and completed discoveries are being communicated to others, both within and outside the domain. Here the emphasis is on rhetorical goals rather than the researcher seeking new ideas. The kinds of analogies that occur in this setting will provide insight into whether long distance analogies occur in the descriptions of research long after it has been completed.

Thus, the lab group/colloquium distinction captures much of the life cycle from ongoing discoveries to final presentations of discoveries, while holding several other features constant (e.g., verbal format and length of presentation). By contrasting the analogies used in each context, we hope to provide some insight into the paradox of long-distance analogies: are they used as sources of ideas by scientists, or do they simple serve rhetorical functions, appearing primarily after the discoveries have been made?

In this paper, we will analyze discussions that occur in lab group and colloquia in various psychology domains. Of particular focus will be the frequency, kind, and use of analogies in each setting. Do different kinds of analogies occur in each setting, and are different kinds of analogies used differently in each setting?

Methods

For this research, the sources of data were a series of video and audio-taped recordings of the proceedings of two psychology lab groups and one colloquium series, all from within one psychology department. The psychology department is part of a major university in a large Eastern US city. The department has over 100 faculty, students, postdocs, and research assistants. Each lab group presentation involved one researcher presenting research that was about to be conducted or had just been conducted but not fully analyzed. The colloquium series presenters were from around the world and were invited as colloquium speakers because of the established excellence of their work.

The data set included three colloquia sessions, which were composed of the primary speaker and 30 to 60 audience members (a mix of faculty, postdocs, graduate students, and research assistants). Two of the speakers were well-known, senior researchers from other universities. The third speaker was a senior, internal faculty member. The colloquia range from 70-90 minutes in length, and topics included social psychology, cognitive psychology, and

Table 1: Coding dimensions and example analogies for each code.

Dimension	Codes	Examples
Source	Within Domain	Other than that, the structure of the study is just like the first one.
	Between Domain	Well, a cognitive resource consumption is really bearing uh the analogy that we use to go in and look at the brain activation. That is, in the model you're using up more and more resources, [that is] you're using up a higher proportion of resources in more difficult tasks.
Use	Used	...there's a kind of linear relationship between the two..., but only when you get up close to the limits do you vastly increase the difficulty, [which is] measured by brain activity. It suggests a somewhat different underlying model. I think disk drive, [with a] disk drive you get similar effects. It's not a linear...
	Mentioned	I'm doing a similar study [to the Siegler and Jenkins study] right now at the ** school
Similarity Base	Similarity	Um, but our production system architecture has some particular properties that are connectionist like... representational elements are conceptualized as having faded levels of activation. They can be more or less active depending on how accessible they are to the system.
	Dissimilarity	And, as you can see, what's different about the forking moves, as opposed to the wins and blocks...
Mode	Asserted	*** All the Above are Asserted Analogies ***
	Anticipated	Well, what about the remember/know paradigm that's gotten so much press lately? Have you tried looking at that?
Analogizer	Presenter	Other than that, the structure of the study is just like the first one.
	Audience	Well, what about the remember/know paradigm that's gotten so much press lately? Have you tried looking at that?

cognitive neuropsychology. None of the participants were aware of the goals of the study.

The five lab group recordings were of the meetings of two different lab groups. One was a developmental psychology lab group comprised of two faculty, three postdocs, three graduate students, and two research assistants. The other was a cognitive psychology lab group comprised of three faculty, six postdocs, five graduate students, and five research assistants. The exact number of participants varied from meeting to meeting. These lab group meetings also were approximately 70-90 minutes long. Both lab groups were run by well-established researchers with research records similar to those of the colloquium speakers.

All speech, including all questions and comments made by audience members were transcribed, producing a total data source of almost 95,000 words. From these transcripts, analogies were identified and coded along five dimensions by two independent coders. Based on the recoding of a subset of the data, there was greater than 90% inter-rater reliability. Differences were easily resolved through discussion.

For the coding, an analogy was defined to occur when existing knowledge from a source object or domain was compared with a target object or domain in order to increase understanding of that target. Any time that a reference was made to the similarity or dissimilarity of two things, in a structural or functional sense, it was counted as an analogy. If something was said to be an attribute of something else or an instance of a category, it was not counted as an analogy. For example, if the statement were to be made that the model "is a connectionist system," it would not be considered an analogy. In addition, if one thing was being simply being re-described using new terminology, it was not considered analogy. For example, if one were to re-describe participant accuracy in terms of d' rather than hit rate, this would not be considered an analogy. Once all the analogies were located and agreed upon, they were coded along several dimensions. These are listed on Table 1, with examples for each.

There were two dimensions that were most critical with respect to our research questions. The first was the "Source" dimension. Here we coded for the distance between source and target according to whether the target came from within the same domain as the source or from a different domain altogether. This dimension was taken from Dunbar (1996) and modified to suit the current domain of psychology. Dunbar distinguished between within organism, between organism, and non-biological analogies. While these distinctions are important and appropriate for molecular biology, they are not applied so easily to psychology. For example, why should analogies between social and cognitive phenomena in humans be less distant than analogies between memory phenomena in rats and humans? Thus, we collapsed all within psychology analogies in a single Within-domain category. Analogies between psychology and other domains (e.g., between psychology and economics) or between two similarly different domains were coded as Between-domain analogies.

The other critical factor was the "Use" dimension: whether or not the analogy was used in reasoning. Analogies could simply be mentioned in the course of the dialogue, just to point out an additional connection that may

or may not be central to the discussion. Such analogies were coded as "Mentioned". In the Mentioned example presented in Table 1, the speaker simply mentioned that they were doing a study analogous to a previous study, without using that analogy to justify a previous point or to motivate a new point. If the speaker made an analogy and used it to illustrate or substantiate a point, or to draw a further inference on the topic at hand, it was coded as "Used". In the Used example presented in Table 1, the speaker brings up the analogy to argue for an alternative conception of cognitive resource consumption. This new conception is then used in the discussion.

Emphasis in previous research has been on similarity analogies as sources of ideas. However, analogies can also be based on dissimilarities between entities. To examine whether such analogies occurred and what their relative frequencies were, we coded for the "Similarity Base" dimension according to whether analogies were based on similarity or dissimilarity.

Another dimension was that of mode. Not all analogies are fully thought through when first proposed. These new, incomplete analogies are interesting because the may track the introduction of new ideas. So, each analogy was coded as "Asserted" if the relation is observed and stated to be present by the analogizer. If the analogizer was asking if a particular relation exists between two systems, or suggesting that one might exist, it was coded as "Anticipated".

A final dimenion of interest is whether the analogies were produced by the presenter or the audience member. This is the "Analogizer" dimension. The ones produced by the audience members are of particular interest because they are potentially bringing new insights to the presenter's topic. Thus, we examined whether between domain analogies were more likely to be given by the presenter (as a rhetorical function) or by the audience member (as a source of ideas).

Results & Discussion

There were a total of 67 analogies identified. Of these 67, 37 came from the 5 lab group transcripts (with between 3 and 15 analogies in each transcript) and 30 from the 3 colloquia transcripts (with between 4 and 15 analogies in each). Thus, analogies did occur with some regularity in each setting, although some presentations in each setting involved more analogies than others.

Because there were more lab group transcripts and because there were roughly a similar number of analogies in each setting, the comparisons between the two settings will focus on proportions rather than absolute frequencies. The proportion of codes on each of the dimension in each setting are presented in Table 2. Since the Ns are relatively low throughout, the statistical tests were done using Fisher Exact tests (which is a very conservative statistic).

For both colloquium and lab group settings, approximately 80% of the analogies were within-domain analogies. This overall high frequency of within-domain analogies support Dunbar's (1996) suggestion that within-domain analogies are the more common occurrence in scientific settings. However, that the proportions did not differ across the two settings was surprising. We had expected a higher proportion of between domain analogies in the colloquium setting.

By contrast, there was a marginally significant difference in the propoportion of Used analogies in each setting. Analogies were almost twice as likely to be used in colloquium setting. Interestingly, the majority of analogies were simply mentioned in both settings.

These overall weak main effects of Setting on Used and Mentioned hid a very important interaction: Used X Mentioned X Setting. Table 3 presents the proportion of analogies used in reasoning separately for within and between domain analogies. Within domain analogies were just as likely to be used in reasoning in colloquia as in lab groups. By contrast, between domain analogies were very likely to be used in reasoning in colloquia but were never used in reasoning in lab groups. Thus, there appears to be good evidence for a difference in how analogies appear to be used in lab groups and colloquia.

To make this effect concrete, consider the following two examples. From the colloquium setting:

> ...striving is real hard to study with the nature of this construct because these are not by definition something that's out there, we achieve it once and you've accomplished it. Like to be a helpful person, you know it's not like, "Well I'll open the door for someone - I'm helpful. I can move on to my next goal." It's kind of like always there, you know, always recurring. So it makes it a little harder to study process questions like...

Here the colloquium speaker is making a between domain analogy (from everyday life to social psychology). Although the analogy is used in reasoning to support his arguments, it is an analogy that was previously thought through and there is no indication that it was used in discovery.

From the lab group setting, an audience member is comparing the results from the current study on children learning addition to a previous study of children learning subtraction:

Table 2: Percent of analogies categorized as within domain, used, similarity based, asserted, and presenter given (separately by setting).

Category	Coll.	Lab Group	p
Within Domain	77%	81%	n.s.
Used	43%	27%	.2
Similarity Based	70%	68%	n.s.
Asserted	93%	97%	n.s.
Presenter	90%	70%	.1

Table 3: Percent of between and within domain analogies used in reasoning (separately by setting).

Type of Analogy	Colloquia	Lab Group	p
Between Domain	71%	0%	.02
Within Domain	35%	33%	n.s.

> Like [with] addition, if you were looking at conception, seems to be more of a conceptual guide, since they don't ever have a procedure, the min strategy, demonstrated for them. Versus in buggy subtraction they're basing it on the fact that they know the correct, they've seen the correct procedure before and that's what they're using. That is a very different kind of thing.

Here, both source and target are from within psychology (both cognitive/mathematical development). Despite the short distance between the source and target, the analogy is an important one, highlighting a potentially crucial factor in the development of new strategies (the topic of the presentation). As one can see by the verbal disfluencies, the analogy is much less well-formed, providing some evidence that it is a new idea. Moreover, the analogy is provided by an audience member rather than by the speaker.

In general, analogies were three times as likely to be generated by an audience member in the lab group setting than in the colloquium setting (although this difference was only marginally significant). However, this did not explain the difference in use of analogies in the two settings: audience members were just as likely to produce between domain analogies as were presenters (21% vs. 21%), and there was no three way interaction with setting. There was also no difference in how likely it was that analogies generated by audience member versus presenter analogies would be used (57% vs. 57%), nor was there a three way interaction with setting.

The frequencies of analogies coded along the Similarity Base dimension follow well with the common understanding of analogies as generally being put in terms of similarity relations. However, there was a significant proportion of dissimilarity based analogies in both settings. There were no differences in the use of similarity vs. dissimilarity based analogies in the two settings.

In both lab group and colloquium settings, almost all of the analogies were asserted rather than anticipated. As one might expect, all three of the anticipated analogies were produced by audience members, and two of those three were in colloquia, where many audience members have relatively little knowledge of the presenter's research topic.

General Discussion

The primary result of this study is that while the relative frequency of analogies held no large differences between presentation settings along any single dimension, the use of analogy did differ significantly as a function of the distance between the source and the target. Specifically, the analogies drawn between domains were only used in the reasoning processes in colloquia. The between domain analogies, while they did occur in lab group settings, were mentioned and not used in reasoning. By contrast, within domain analogies were used in both settings.

These findings support one resolution of the paradox between psychology lab findings and history and philosophy of science claims: long distance analogies appear to be rarely used in the actual discovery processes and, instead, appear to be rhetorical additions added after the fact. Of course, we do not wish to claim that long distance analogies are never used in science. Rather, this study simply suggests: 1) that long distance analogies do not occur frequently as sources of new ideas in scientific settings; 2) within domain analogies serve a much larger role in these settings; and 3) the historical record may be misleading with respect to the role that long distance analogies played in previous discoveries.

Our findings have both some important similarities and differences between those found by Dunbar (1996). Like Dunbar, we found that long-distance analogies are not used in reasoning in lab group settings. This similarity across research groups from very different sciences supports the generality of these findings. However, unlike Dunbar, we did find many (just mentioned) long-distance analogies in the lab group settings. It remains to be seen whether these are differences between psychology and biology: research problems in psychology may be easier to relate to everyday life experiences. Alternatively, it is possible that differences were due to differences in coding analogy: Dunbar's definition of analogy included that it must be used in reasoning, and thus there may have been many mentioned long-distance analogies that were excluded.

If the results of the current study and Dunbar's study are correct, then the consequences for historical studies are quite severe: writings or oral reports made after the discovery may not be trusted regarding the role that analogies played in the discoveries. Of course, psychologists have been long suspicious of the accuracy of retrospective reports (Nisbett & Wilson, 1977; Ericsson & Simon, 1993). Moreover, people have been shown to be very inaccurate regarding the origins of their scientific hypotheses (Dunbar, 1996; Schunn & Dunbar, 1996). Not all historical analyses fall prey to this problem equally: lab books will continue to be a good source of data. Unfortunately, not all scientists are good about keeping organized and complete lab books.

It is also important to note that historical and philosophical analyses will always play an important role in the cognitive science of science, regardless of the accuracy of the current study's results. At the very least, they provide descriptions of important macro-level phenomena in the history of science that need to be further explained (e.g., Boden, 1983; Gentner et al., 1997; Nersessian, 1985; Schunn, Crowley, & Okada, 1998; Thagard, 1988).

The methodology used here represents clear tradeoffs in experimental design: as we increase external validity we lose the benefits of studies conducted in the psychology lab. A few particular issues deserve mention. First, there were only two lab groups and only 8 total sessions across both settings. This issue of small Ns is a common problem with detailed process analyses of real world cognition. However, despite the small N, the key findings involved very large effects that were statistically reliable using conservative statistical tests. To examine how stable the findings are within each setting, we are currently extending the analyses to a larger set of colloquia and lab group presentations.

Second, it is possible that important new ideas are developed or discoveries are made outside of the lab group setting and are not discussed until much later presentations (e.g., a colloquium talk). However, the five lab group presentations were of research at various phases of progress (from just started to almost completed), and in no case was there use of a between domain analogy in reasoning. It is

unclear why important between domain analogies would remain hidden in each of these cases.

Another potential problem is that there were different topics in the different settings. For example, the topic of one of the colloquia was social psychology, whereas none of the lab groups involved social psychology. However, there was a variety of topics in both settings, with some overlap, and the main findings appeared consistent across topics within settings. Thus, it is likely that these findings will generalize across other topics as well.

It is also true that there were different people present in each of the settings, both as presenters and audience members. More crucially, some of the lab group presenters were graduate students and postdocs—a very different level from the colloquium presenters. However, all of the research presented in the lab groups was designed and conducted in close consultation with the primary faculty researcher. Moreover, there was consistency in the main findings across the different presentations, suggesting that the main results generalize across different presenter levels.

Finally, to anticipate a suggestion, it is clear that a longitudinal design should be the next step. Fortunately, the lab groups that appeared in this study were quite strong, and it is likely that the results from their studies will appear in the literature in the near future. Observation of more formal presentations of these same projects would be a logical follow-up project, and the current study suggests that it will be a worthwhile undertaking.

Acknowledgments

We would like to thank Greg Trafton, Erik Altmann, Susan Trickett, Sheryl Miller, and Audrey Lipps for comments made on earlier drafts of this paper. Work on this paper was supported by a grant from the Mitsubishi Bank Foundation.

References

Blanchette, I. & Dunbar, K. (1997). Constraints Underlying Analogy Use in a Real-World Context: Politics. Poster presented at the *19th Annual Meeting of the Cognitive Science Society*, Stanford, CA.

Boden, M. (1993). *The creative mind: Myths & mechanisms*. New York, NY: BasicBooks, Inc.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J. Sternberg & J. Davidson (Eds.), *Mechanisms of insight*. Cambridge MA: MIT press. pp 365-395.

Dunbar, K. (1996). How scientists think: Online creativity and conceptual change in science. In T.B. Ward, S.M. Smith, & S.Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery and Change*. Washington, DC: APA Press.

Dunbar, K., & Baker, L. M. (1994). Goals, analogy, and the social constraints of scientific discovery. *Behavioral and Brain Sciences, 17*, 538-539.

Ericsson, K. A., & Simon H. A. (1993). *Protocol analysis: Verbal reports as data* (Revised Edition). Cambridge, MA: MIT Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gentner, D., Brem, S., Ferguson, R. W., Markman, A., Levidow, B. B., Wolff, P., & Forbus, K. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *Journal of the Learning Sciences, 6*(1), 3-40.

Gentner, D., & Jeziorski, M. (1989). Historical shifts in the use of analogy in science. In W. R. S. B. Gholson, Jr., R. A. Beimeyer, & A. Houts (Eds.), *The psychology of science: Contributions to metascience* (pp. 296-325): Cambridge University Press.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 19) (pp. 59-87). New York: Academic Press.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*, 332-340.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Koestler, A. (1964). *The act of creation*. London: Macmillian.

Mandler, J. M., & Orlich, F. (1993). Analogical transfer: The roles of schema abstraction and awareness. *Bulletin of the Psychonomic Society, 31*(5), 485-487.

Nersessian, N. J. (1985). Faraday's field concept. In D. Gooding & F. James (Eds.), *Faraday rediscovered* (pp. 175-187). London: Macmillan.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*(3), 510-520.

Rhodes, R. (1986). *The making of the atomic bomb*. New York: Simon & Schuster.

Ross, B. H. (1989). Distinguishing types of superficial similarities: different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 456-468.

Schunn, C. D., & Dunbar, K. (1996). Priming, analogy, and awareness in complex reasoning. *Memory & Cognition, 24*(3), 271-284.

Schunn, C. D., Crowley, K., & Okada, T. (1998). The growth of multidisciplinarity in the Cognitive Science Society. *Cognitive Science*, 22(1), 107-130.

Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press.

Ueda, K. (1997). Actual use of analogies in remarkable scientific discoveries. Paper presented at the *19th Annual Meeting of the Cognitive Science Society*, Stanford, CA.

VanLehn, K (1998). Analogy events: How examples are used during problem solving. *Cognitive Science. 22*(3). 347-388

Young, K.M., & Leinhardt, G. (in press) *Wildflowers, sheep, and democracy: The role of analogy in the teaching and learning of history*.