International Journal for the Scholarship of Teaching and Learning

Volume 17 | Number 1 Article 12

May 2023

A Really Good Example Helps Learning About an Abstract Concept

Ava Funkhouser

Elena Nicoladis
University of British Columbia, elena.nicoladis@ubc.ca

A Really Good Example Helps Learning About an Abstract Concept

Abstract

University students are often asked to learn abstract concepts. Abstract concepts are hard to learn. Giving specific examples can help learning abstract concepts. These examples might limit understanding to the similarities between the abstract domain and particular examples. The primary purpose of this study was to test whether exposure to multiple examples would lead to better learning than exposure to a single example. Secondarily, we were interested in whether there was any particularly effective example. Introductory psychology students were invited to learn about the abstract concept of semiotics, through either 1) three of five distinct examples or 2) a single example presented three times. We assessed learning through definitions, transfer to a novel example, and self-report. The results showed no support for the hypothesis that exposure to multiple examples led to better learning. There was, however, one particular example that was more memorable and resulted in better learning. These results have implications about how best to teach abstract concepts.

Keywords

examples; abstract concepts; multiple examples; abstraction; semiotics

A Really Good Example Helps Learning About an Abstract Concept

Ava Funkhouser and Elena Nicoladis

University of British Columbia

Received: 30 June 2021; Accepted: 24 September 2021

University students are often asked to learn abstract concepts. Abstract concepts are hard to learn. Giving specific examples can help learning abstract concepts. These examples might limit understanding to the similarities between the abstract domain and particular examples. The primary purpose of this study was to test whether exposure to multiple examples would lead to better learning than exposure to a single example. Secondarily, we were interested in whether there was any particularly effective example. Introductory psychology students were invited to learn about the abstract concept of semiotics, through either 1) three of five distinct examples or 2) a single example presented three times. We assessed learning through definitions, transfer to a novel example, and self-report. The results showed no support for the hypothesis that exposure to multiple examples led to better learning. There was, however, one particular example that was more memorable and resulted in better learning. These results have implications about how best to teach abstract concepts.

University students are often called upon to learn abstract concepts, concepts with referents that cannot be detected through the senses, like *anthropomorphism* or *aerodynamics.* For example, students in literature courses may be asked to define anthropomorphism or identify instances of anthropomorphism in assigned texts. Learning abstract concepts is difficult (Borghi et al., 2017). The most obvious difficulty in learning abstract concepts is that, by definition, there is no referent for these concepts that is available to the senses (Borghi et al., 2017). One approach that helps learners with abstract concepts is to explain the connection between abstract concepts and the physical world, with examples, metaphors, analogies, computer simulations, and/or images (Donnelly & McDaniel, 1993; Duit, 1991; Evans & Evans, 1989; Falloon, 2019; Gentner & Asmuth, 2019; Halpern et al., 1990; Morgan & Reichert, 1999; Orgill & Bodner, 2007, 2004; Podolefsky & Finkelstein, 2007; Singh, 2010; Wormeli, 2002).

In the present study, we focus on the role of examples in learning abstract concepts. Studies have shown that students spontaneously generate examples when studying material based on abstract concepts (Gurung et al., 2010). However, students do not always generate valid or useful examples of abstract concepts (Zamary et al., 2016). Instructor-provided examples, however, have been shown to help learning and retention, both immediately and over a few days (Balch, 2005; Rawson & Dunlosky, 2016; Rawson et al., 2015; Zamary & Rawson, 2018). One way in which examples can support learning is by providing a mnemonic, allowing learners to access their knowledge quickly and easily (Zamary & Rawson, 2018; see also Orgill & Bodner, 2007, for similar results with metaphors). Moreover, by understanding an abstract concept through the physical world, learners can understand the rules or principles that constrain the abstract concept (Evans & Evans, 1989; Halpern et al., 1990; Jeppsson, Haglund, Amin, & Strömdahl, 2013; Singh, 2010).

There are, however, some potential pitfalls in learning abstract concepts through examples. Any particular example does not capture the entirety of the abstract concept. It is possible that learning an abstract concept based on a single concrete example could lead to misunderstanding. Research with learning abstract concepts based on metaphors has shown that learners sometimes do not generalize beyond the specificities of a particular metaphor when thinking about an abstract concept (Evans & Evans, 1989; Falloon, 2019; Jamrozik et al., 2016; Morgan, 1997; Podolefsky & Finkelstein, 2007; Spiro et al., 1989; see similar reasoning in McDaniel et al., 2018).

One way to address this potential limitation is to provide learners with multiple examples. By hearing multiple examples, learners could recognize the similarities across the examples and generalize to the characteristics of the abstract concept. While we know of no research testing this approach with examples, exposing learners to multiple metaphors leads to deeper understanding than exposing learners to one metaphor (Arzarello, 2006; Harrison & De Jong, 2005; Mildenhall & Sherriff, 2016; Ott, Brünken, Vogel & Malone, 2018; Spiro et al., 1989). Moreover, learners with exposure to multiple metaphors make errors related to a specific metaphor relative to learners with exposure to only one metaphor (Ott et al., 2018; Spiro et al., 1989). The primary purpose of this study was to test whether these results generalize to examples. If so, then learning an abstract concept through exposure to multiple examples should be better than learning an abstract concept through exposure to a single example.

A secondary purpose of this study was to test whether there were particular examples that were better than others to support learning. In a study on analogies, biochemistry students reported that effective analogies were simple, easy to remember, and were related to highly familiar concrete concepts (Orgill & Bodner, 2004). Note that our prediction that some examples will be better than others is not mutually exclusive from the prediction that multiple examples will be helpful. Indeed, it is possible that exposure to multiple examples is highly effective because it increases the likelihood of exposure to a particularly useful example (see Branigan et al., 2012; Evans & Evans, 1989, for similar reasoning).

ASSESSING LEARNING

In this study, we assessed learning with three different measures: definitions, transfer to novel examples, and a self-report measure of learning.

Definitions are often used to assess learners' understanding of abstract concepts (Gablasova, 2015; Johnson & Anglin, 1995; Rawson et al., 2015; Walker, 2001) because definitions reveal both what students have understood and what they have misunderstood. In order to define an abstract concept when presented with examples, learners must find the generalities among the specifics to understand the underlying rules or principles (Johnson & Anglin, 1995). This same kind of generalization likely underlies learners' ability to transfer their understanding of a presented example to recognize a novel valid example of the same concept (Goldwater et al., 2018; Knoop-van Campen et al., 2020). We therefore predicted that these two objective measures of learning would be positively correlated.

Many studies have found no relationship between objective and subjective measures of learning (Potter, 2013), often because learners are overly confident of their own knowledge (Zamary et al., 2016). We hypothesized that self-reported learning might unrelated to the objective measures of learning.

This study

In this study, we taught introductory psychology students about semiotics. We expected them to have little knowledge about semiotics because it is not a topic typically covered in the high school curriculum.

The primary purpose of this study was to test whether presenting learners with multiple examples leads to better learning than presenting them with only one example. If so, then learners should be better at defining an abstract concept and transferring to a novel example after hearing multiple examples than after a single example.

A secondary purpose of this study was to test for particularly effective examples. We operationalized the effectiveness of examples in terms of both memorability (i.e., the example used when explaining the abstract concept to a friend) and learning (i.e., better definitions, better transfer to novel examples, and higher self-reports of understanding). We hypothesized that the memorability would be correlated with learning success.

METHOD

Participants

132 participants were included in the final sample. All participants were introductory psychology students, averaging 19.0 years of age (SD = 1.6) and mostly female (91 females, 41 males, 1 non-binary). We asked participants to rate their degree of familiarity before the study about semiotics on a 10-point scale (with 1 being "extremely unfamiliar" and 10 "extremely familiar"). As expected, the participants reported little familiarity with semiotics (M = 1.53, SD = 2.26).

Materials

Participants listened to audio recordings of the examples. These recordings feature a male describing an example of a semiotic system and range between 45 seconds to 2 minutes. The speaker is an English professor and well versed in semiotics. The professor was told ahead of time that we would ask him about semiotics but did not specifically mention ahead of time that we would be asking for examples. When obtaining the recordings, the researcher asked him to speak candidly. He had little time to prepare in order to closely mimic spontaneous speech that might occur in the context of a lecture. We reasoned that spontaneous speech would be more engaging than prepared speech. Once he had given all the examples, he agreed that all of the examples were valid examples.

There were a total of five examples (See Appendix A). The examples used were Morse code (i.e., a meaningless sound which only gains meaning within a system), Fashion (i.e., clothes can signal status and occupation, but only within a system), Traffic lights (i.e., red lights have no inherent meaning, but gain meaning within the system of traffic lights), Classroom (i.e., desks and podiums have no inherent meaning, but within a classroom signifies roles), and Chess board (i.e., pawns and queens have no inherent meaning, but gain meaning within the system of the game of chess).

We used three measures to assess learning: definition, transfer to novel valid/invalid examples, and self-reported learning. For the definitions, the instructions were: "Define semiotics in your own words. Please be specific, try to include as much detail as possible." For the transfer to a novel example, the participants were asked to rate novel example on a scale of 1 (extremely inaccurate) to 10 (extremely accurate). One of the examples was valid and the other invalid (see Appendix B), with both examples using similar language to the original examples. Finally, after participants heard the example (or examples), we asked them to rate their understanding on a 10-point scale (with 1 being "extremely poor" and 10 "extremely good"). In order to measure learning, we subtracted their self-rating of their prior knowledge from the present knowledge. This difference constituted the measure of self-reported learning.

To assess the memorability of particular examples, we asked participants to say how they would explain semiotics/equilibrium to a friend.

PROCEDURE

Participants were randomly assigned to a condition: single example (SE) or multiple examples (ME). We assigned more participants to the SE condition (N = 76) than to the ME condition (N = 56), in order to test for possible effects of particular examples.

Within the ME condition, the three examples were presently in random order for each participant. Within the SE condition, the example was selected randomly for each participant and repeated three times. In the SE condition, 14 participants heard the Traffic light analogy, 17 the Fashion, 15 Morse code, 18 the Chess board, and 12 the Classroom. A research assistant showed the participants into a cubicle and then fitted them with headphones. The study was presented via a Qualtrics survey. The participants were asked to define semiotics, rate their prior and present knowledge about the concept, and evaluate the valid and invalid novel examples. Finally, the participants received a debriefing form and had an opportunity to ask questions.

Coding

The quality of the definitions was rated on a scale from 0-3, with half-points assigned. The definitions had to include three separate parts, for one point each:

1. A system of meaning and communication
2. There is a relationship that has no inherent meaning
3. The relationship gains meaning within the system.

Here is an example of a 3-point answer from a participant in the single perspective condition: "Semiotics refers to the relationship between certain words or signals and their meanings." (1 point) "These relationships are arbitrary, and on their own would not be related to one another" (1 point) "but within a specific closed system there is a clear connection between the two." (1 point). A 0 point answer was "The connection between teacher and student," and a 1.5 point answer was "Semiotics is the idea that words or figures used to represent or describe a concept can have meaning in that specific closed system only. Based on semiotics, a figure can only gain value when it is used in a specific system of figures."

To test on the inter-rater reliability of this coding, 20% of the data (i.e., 26 participants' definitions) was randomly chosen for independent coding by a second coder. For these 26 participants, the second coder averaged 0.85 (SD = 0.83) and the first coder 0.79 (SD = 0.79). There was no significant difference between the two reviewers' coding on an independent-samples t-test, t (50) = 0.26, p = .80. Moreover, the two reviewers' scores were highly correlated, r (24) = 0.83, p < .00001. For the analyses, we retained the first coder's scores.

To measure the memorability of particular analogies, we counted the number of participants in the ME condition who used a particular example to explain the abstract concept for a friend. To test if memorability was correlated with learning, we correlated the memorability (from the ME condition) with the learning measures for the participants in the SE condition.

RESULTS

Definitions

We predicted that exposure to multiple examples (ME) would lead to better definitions than exposure to single examples (SE). Contrary to this prediction, there was no significant difference between the definition scores in the ME condition (M = 0.97, SD = 0.81) and those in the SE condition (M = 1.04, SD = 0.86), t (130) = 0.45, p = .66.

The lack of difference between the two conditions could have been due to the restricted scale we used to rate definitions. To test that possibility, we counted the number and percentage of participants who achieved each definition score (summarized in Table 1). We then tested if there were any differences between conditions on the numbers of participants assigned each score using a chi-square test. This analysis also revealed no significant difference between conditions, c^2 (df = 6, N = 132) = 8.62, p = .20.

In sum, we found no evidence supporting the prediction that hearing multiple examples would lead to better definitions than hearing single examples.

Table 1. Percentage (n) participants receiving each definition score

	Multiple examples	Single example
0	25.4% (14)	13.4% (9)
0.5	14.6% (8)	25.4% (17)
1	27.3% (15)	29.9% (20)
1.5	16.4% (9)	6.0% (4)
2	9.1% (5)	11.9% (8)
2.5	3.6% (2)	6.0% (4)
3	3.6% (2)	7.5% (5)

Transfer to novel analogies by condition

We predicted that exposure to ME would lead to better transfer to novel examples than exposure to SE. We compared the participants' ratings of the novel valid example with the novel invalid example by condition with a 2 x 2 [Valid/Invalid x Condition] ANOVA, with Valid/Invalid as a repeated measure. This analysis showed a main effect for Valid/Invalid, F (1, 129) = 80.16, p < .001, η^2 = .383, but no main effect of Condition, F (1, 129) = 0.10, p = .75, η^2 = .001, and no interaction, F (1, 129) = 0.03, p = .86, η^2 < .001. Table 2 summarizes the results. We found no evidence supporting our prediction that exposure to ME leads to greater transfer to novel analogies than exposure to SE.

Table 2. Average (SD) for rating novel valid and invalid analogies

	Multiple examples	Single example	Total
Valid	7.93 (1.96)	7.79 (1.89)	7.86 (.17)
Invalid	5.30 (2.42)	5.27 (2.65)	5.29 (.23)

Note: Validity ratings could range from 1 (extremely inaccurate) to 10 (extremely accurate)

Self-report of learning by condition

Our third prediction was that exposure to ME would lead to higher self-reported learning than exposure to SE. A 2 x 2 [Before/After x Condition] ANOVA, with Before/After as a repeated measure showed a main effect for Before/After, F (1, 130) = 322.61, p < .001, η^2 = .713, but no main effect of Condition, F (1, 130) = 0.37, p = .54, η^2 = .003, and no interaction, F (1, 130) = 0.04, p = .84, η^2 < .001. Table 3 summarizes the results. There was no evidence supporting our prediction that exposure to ME leads to greater self-reported learning than exposure to SE.

Table 3. Average (SD) for self-reported knowledge before and after the study

	Multiple examples	Single example	Total
Before	1.61 (2.27)	1.46 (2.27)	1.52 (2.26)
After	5.75 (2.13)	5.51 (2.15)	5.61 (2.13)

Note: Knowledge ratings could range from 1 (extremely poor) to 10 (extremely good)

Correlations between measures of learning

Table 4 summarizes the correlations between, definitions, ratings of novel valid/invalid examples, and self-report of learning. As can be seen in that Table, the better definitions the participants gave of semiotics, the more they endorsed both the valid and invalid novel examples. The more they endorsed a novel valid example of semiotics, the higher their self-reported learning. These results do not strongly support our prediction that the objective measures would be highly inter-correlated and the subjective measure uncorrelated with the objective measures.

Table 4. Correlations between measures of learning

	1. Definition	2. Valid	3. Invalid
1. Definition	-		
2. Valid	.359**	-	
3. Invalid	.392**	-.037	-
4. Self-report	.144	.218*	-.050

* p < .05; ** p < .01

Memorability

Table 5 summarizes the data for each of the examples. We first tested whether there was a particularly effective example. Within the SE condition, a one-way ANOVA on the definition scores demonstrated that there was a significant main effect for example, F (4, 71) = 8.11, p < .001, η^2 = .314. LSD post-hoc tests revealed that the Traffic lights analogy was better than all the other analogies and the Morse code analogy was better than the Classroom

Table 5. Summary by analogy of memorability and average (SD) measures of learning					
	Memorable†	**Definitions**	**Valid**	**Invalid**	**Self-reported learning**
Traffic lights	28	1.93 (1.00)	7.00 (2.25)	2.00 (2.69)	4.07 (2.13)
Morse code	4	1.23 (.72)	4.80 (2.27)	1.07 (1.87)	4.67 (2.50)
Fashion	9	.88 (.49)	5.06 (2.96)	1.00 (1.70)	4.47 (1.62)
Chess board	4	.67 (.71)	4.83 (2.79)	2.33 (2.99)	3.44 (2.94)
Classroom	3	.54 (.72)	4.73 (2.45)	0.67 (1.15)	3.58 (2.94)

† Number of participants in the ME condition who used the example in explaining the concept to a friend; numbers do not add up to 56 because some participants gave novel examples

and the Chessboard. As for the transfer to novel examples, there was a difference between the Valid and the Invalid novel analogies, $F(1, 70) = 39.00$, $p < .001$, $\eta^2 = .358$, and a significant difference by Example, $F(4, 70) = 3.83$, $p = .007$, $\eta^2 = .180$. There was no interaction between Analogy and Valid/Invalid, $F < 1$, *ns*, $\eta^2_p = .020$. As can be seen in Table 4, the effect of Example is that Traffic lights is higher than everything else. For self-reported learning, there was no difference by example, $F < 1$, *ns*, $\eta^2_p = .042$. Spearman rank-order correlations showed that memorability was highly positively correlated to average definition scores across the five examples, *rho* = .821, $p = .09$, and to ratings of the novel valid example, *rho* = .975, $p = .005$, but not to the ratings of the novel invalid example, *rho* = .410, $p = .49$, or the self-reported learning, *rho* = .308, $p = .61$.

In sum, in learning about semiotics, there was a clear most effective example: Traffic lights. It was the most memorable, resulted in the highest definition scores, and led to the highest acceptance of a novel valid concept as well as a high rejection of a novel invalid concept. Furthermore, among the five examples, memorability was positively related to how well participants defined semiotics and how highly they accepted novel valid concepts.

DISCUSSION

The primary purpose of this study was to test whether exposure to multiple examples would lead to better learning than exposure to a single example when learning about the abstract concept of semiotics. Previous research has shown that exposure to multiple analogies supports a deeper learning of abstract concepts (Ainsworth et al., 2002; Arzarello, 2006; Confrey & Smith, 1994; Mildenhall & Sherriff, 2016; Ott et al., 2018; Spiro et al., 1989). We found no support for our prediction. Participants who heard only one example of semiotics and participants who heard three different examples performed equivalently on definitions, transfer to novel examples, and self-reported learning.

One possible reason for finding no difference between multiple examples and a single examples is that the learners in this study only gained very shallow knowledge of semiotics, regardless of condition. The participants' average definition score was about one out of three, suggesting that their definitions were sufficiently describing one of the three aspects of the full definition. Similarly, while the participants gave higher endorsements of a novel valid example than a novel invalid example, their ratings of the invalid example were high (averaging about 5 out of 10). They also did not rate their own knowledge of semiotics very highly even after the study (averaging about 5.5 on a scale of 10). It is possible that more extensive exposure to an abstract concept is necessary for the number of examples to make a difference. Consistent with this argument, one study focused a high school teacher used multiple analogies in teaching the concept of chemical equilibrium over the course of several months (Harrison & De Jong, 2005). His students subsequently demonstrated deep and complex knowledge about equilibrium, including aspects of equilibrium that are challenging for university students. Future studies can be designed to test for possible advantages of presenting multiple examples of abstract concepts in the context of longer-term learning.

A secondary purpose of the present study was to test whether there were some particularly effective examples (see Evans & Evans, 1989; Orgill & Bodner, 2004). There was clear evidence that one example was particularly effective for learning semiotics in this study. The Traffic light analogy was highly memorable, led to high-definition scores, high acceptance of a valid novel concept, and high rejection of an invalid novel concept. It was not entirely clear why that example was so effective. For analogies, Orgill and Bodner (2004) identified three characteristics that they thought would be highly related to effectiveness: simplicity, ease of remembering, and familiarity. While traffic lights may have been more familiar to many participants than Morse code or chess, it is not clear that the example of traffic lights was simpler or more familiar than the examples related to fashion or classrooms. One possibility is that many of the participants could have had experience with cross-cultural differences in traffic lights. For example, in France, yellow lights are labelled orange and in Korea, green lights are labelled blue. We did not collect data about the participants' experience with different cultures (or any of the relevant domains for the specific examples), but 48% of the participants spoke a language other than English as their first language. Moreover, the university where the research was carried out attracts many international students. Note that this university requires a high level of English proficiency in order to be admitted so it is unlikely that English proficiency affected the results. Our explanation that many of our participants might have had some familiarity with cross-cultural differences is highly speculative and future research can address why some examples are better than others in learning.

One important point to keep in mind about the present study is that we considered the participants as a group. Some studies have shown individual differences in learning through exemplars, with some students learning better through exemplars and others through abstract rules (Foster, Rawson, & Dunlosky, 2018; McDaniel, Cahill, Frey, Rauch, Doele, Ruvolo, & Daschbach, 2018; McDaniel, Cahill, Robbins, & Wiener, 2014). These individual differences have long-term implications for learning at university (Frey, Cahill, & McDaniel, 2017). Future studies could include measures of those individual differences in learning when testing how metaphors help learning abstract concepts (see also Bjork, Dunlosky, & Kornell, 2013, for further discussion on individual differences among learners).

CONCLUSION

To conclude, this study has shown that a good example can allow learners to gain rapid insight into a novel abstract concept. Further research is needed to identify what makes particular examples highly effective. These results add to the literature showing that teaching abstract concepts by using examples can support learning about abstract concepts (Balch, 2005; Rawson & Dunlosky, 2016; Rawson et al., 2015; Zamary & Rawson, 2018). However, instructors might wish to pilot their examples before using them in the classroom (see Orgill & Bodnerm, 2004, for a similar suggestion for metaphors).

CONTACT

Elena Nicoladis <elena.nicoladis@ubc.ca>

REFERENCES

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *The Journal of the Learning Sciences*, *11*, 25–61.

Arzarello, F. (2002). Semiosis as a multimodal process. *Revista Latinoamericana de Investigación En Matemática Educativa, 9,* 267-300.

Balch, W. R. (2005). Elaborations of introductory psychology terms: Effects on test performance and subjective ratings. *Teaching of Psychology, 32*, 29–34.

Bergquist, W. & Heikkinen, H. (1990). Student ideas regarding chemical equilibrium: What written test answers do not reveal. *Journal of Chemical Education*, *67*, 1000-1003.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444.

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, *143*(3), 263-292.

Branigan, H. P., Catchpole, C. M., & Pickering, M. J. (2011). What makes dialogues easy to understand? *Language and Cognitive Processes*, *26*, 1667–1686.

Confrey, J., & Smith, E. (1994). Exponential functions, rates of change, and the multiplicative unit. *Educational Studies in Mathematics*, *26*, 135–164. https://doi.org/10.1007/BF01273661

Donnelly, C. M. & McDaniel, M. A. (1993). Use of analogy in learning scientific concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 975-987

Duit, R. (1991). On the role of analogies and metaphors in learning science. *Science Education*, *75*(6), 649-672.

Evans, R. D. & Evans, G. E. (1989). Cognitive mechanisms in learning from metaphors. *The Journal of Experimental Education*, *58*(1), 5-19.

Falloon, G. (2019). Using simulations to teach young students science concepts: An Experiential Learning theoretical analysis. *Computers & Education*, *135*, 138-159.

Foster, N. L., Rawson, K. A., & Dunlosky, J. (2018). Self-regulated learning of principle-based concepts: Do students prefer worked examples, faded examples, or problem solving?. *Learning and Instruction*, *55*, 124-138.

Frey, R. F., Cahill, M. J., & McDaniel, M. A. (2017). Students' concept-building approaches: A novel predictor of success in chemistry courses. *Journal of Chemical Education*, *94*(9), 1185-1194.

Gablasova, D. (2015). Learning technical words through L1 and L2: Completeness and accuracy of word meanings. *English for Specific Purposes*, *39*, 62-74.

Gentner, D. & Asmuth, J. (2019). Metaphoric extension, relational categories, and abstraction. *Language, Cognition and Neuroscience*, *34*(10), 1298-1307.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001) Metaphor is like analogy. In Gentner, D., Holyoak, K.J., & Kokinov, B.N. (Eds.) *The Analogical mind: Perspectives from cognitive science* (pp. 199-255). Cambridge, MA: The MIT Press.

Goldwater, M. B., Don, H. J., Krusche, M. J. F., & Livesey, E. J. (2018). Relational discovery in category learning. *Journal of Experimental Psychology: General, 147*, 1–35.

Gorodetsky, M. & Gussarsky, E. (1986). Misconceptualization of the chemical equilibrium concept as revealed by different evaluation methods. *European Journal of Science Education*, *8*(4), 427-441.

Gurung, R.A.R., Weidert, J., & Jeske A. (2010). Focusing on how students study. *Journal of the Scholarship of Teaching and Learning, 10*(1), 28-35.

Halpern, D. F., Hansen, C., & Riefer, D. (1990). Analogies as an aid to understanding and memory. *Journal of Educational Psychology*, *82*(2), 298-305.

Harrison, A. G., & De Jong, O. (2005). Exploring the use of multiple analogical models when teaching and learning chemical equilibrium. *Journal of Research in Science Teaching*, *42*(10), 1135-1159.

Jamrozik, A., McQuire, M., Cardillo, E. R., & Chatterjee, A. (2016). Metaphor: Bridging embodiment to abstraction. *Psychonomic Bulletin and Review*, *23*(4), 1080-1089.

Jeppsson, F., Haglund, J., Amin, T. G., & Strömdahl, H. (2013). Exploring the use of conceptual metaphors in solving problems on entropy. *Journal of the Learning Sciences*, *22*(1), 70-120.

Johnson, C. J. & Anglin, J. M. (1995). Qualitative developments in the content and form of children's definitions. *Journal of Speech, Language, and Hearing Research*, *38*(3), 612-629.

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2020). Effects of audio support on multimedia learning processes and outcomes in students with dyslexia. *Computers & Education*, *150*, 103858. https://doi.org/10.1016/j.compedu.2020.103858

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

McDaniel, M. A., Cahill, M. J., Frey, R. F., Rauch, M., Doele, J., Ruvolo, D., & Daschbach, M. M. (2018). Individual differences in learning exemplars versus abstracting rules: Associations with exam performance in college science. *Journal of Applied Research in Memory and Cognition*, *7*, 241-251.

McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: Stable tendencies for learning exemplars versus abstracting rules. *Journal of Experimental Psychology: General, 143*, 668–693.

Mildenhall, P., & Sherriff, B. (2018). Using multiple metaphors and multimodalities as a semiotic resource when teaching year 2 students computational strategies. *Mathematics Education Research Journal*, *30*, 383–406. https://doi.org/10.1007/s13394-017-0212-8

Morgan, G. (1997) *Images of organization*. Thousand Oaks CA: Sage

Morgan, S. E. & Reichert, T. (1999). The message is in the metaphor: Assessing the comprehension of metaphors in advertisements. *Journal of Advertising*, *28*(4), 1-12.

Oberlechner, T., Slunecko, T., & Kronberger, N. (2004). Surfing the money tides: Understanding the foreign exchange market through metaphors. *British Journal of Social Psychology*, *43*(1), 133-156.

Orgill, M. & Bodner, G. (2007). Locks and keys: An analysis of biochemistry students' use of analogies. *Biochemistry and Molecular Biology Education*, *35*(4), 244-254.

Orgill, M. & Bodner, G. (2004). What research tells us about using analogies to teach chemistry. *Chemistry Education Research and Practice*, *5*(1), 15-32.

Ott, N., Brünken, R., Vogel, M., & Malone, S. (2018). Multiple symbolic representations: The combination of formula and text supports problem solving in the mathematical field of propositional logic. *Learning and Instruction*, *58*, 88–105. https://doi.org/10.1016/j.learninstruc.2018.04.010

Palmer, S.E. (1978) Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd, (Eds.) *Cognition and Categorization* (pp. 259-303). Hillsdale, NJ: Lawrence Erlbaum Associates.

Piquette, J. S. & Heikkinen, H. W. (2005). Strategies reported used by instructors to address student alternate conceptions in chemical equilibrium. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *42*(10), 1112-1134.

Podolefsky, N. S. & Finkelstein, N. D. (2007). Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies. *Physical Review Special Topics-Physics Education Research*, *3*(2), 020104.

Porter, S. R. (2013). Self-reported learning gains: A theory and test of college student survey response. *Research in Higher Education*, *54*, 201-226.

Rawson, K. A. & Dunlosky, J. (2016). How effective is example generation for learning declarative concepts? *Educational Psychology Review, 28*(3), 649–672.

Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review, 27*, 483–504.

Singh, K. (2010). Metaphor as a tool in educational leadership classrooms. *Management in Education*, *24*(3), 127-131.

Spiro, R.J., Feltovich, P.J., Coulson, R.L., & Anderson, D. K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconception in advanced knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 498-531). Cambridge, MA: Cambridge University Press.

Thibodeau, P. H. & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLOS One*, *6*, e16782. https://doi.org/10.1371/journal.pone.0016782

Walker, S. J. (2001). Cognitive, linguistic, and social aspects of adults' noun definitions. *Journal of Psycholinguistic Research*, *30*(2), 147-161.

Wormeli, R. (2002). *Metaphors & analogies: Power tools for teaching any subject*. Portland, ME: Stenhouse Publishers.

Zamary, A. & Rawson, K. A. (2018). Which technique is most effective for learning declarative concepts—Provided examples, generated examples, or both? *Educational Psychology Review*, *30*(1), 275-301.

Zamary, A., Rawson, K. A., & Dunlosky, J. (2016). How accurately can students evaluate the quality of self-generated examples of declarative concepts? Not well, and feedback does not help. *Learning and Instruction*, *46*, 12-20.

APPENDIX A

EXAMPLES FOR SEMIOTICS

Introduction Definition:
"So semiotics is the study of communication systems such as language. There's two crucial interventions of modern semiotics. The first is the idea that the relationship between a w– concept and the word that we use to represent that concept is absolutely arbitrary. The second, semiotics argues that the meaning of any given word only makes sense within a closed system."

Example 1: *Morse Code*
"One way to look at this would be through the example of morse code. A system of meaning in which "dot dash" means "A" and "dash dot" means N. There's nothing inherent to the "dot dash" that should mean A, however it is only within that particular system that that, uh, figure comes to hold a certain kind of value. We know that " dosh-- that dot dash" means A because it's not "dash dot", N. It's only in that system of relationship of differences through which the "dot dash" gains meaning."

Example 2: *Fashion*
"One way to look at this would be fashion, or clothing, itself which can be understood as a language or a kind of communication system. When I go to work, as a professor, I wear a jacket or a suit and I use this to try to c-- produce the meaning that I hold a certain kind of role in a professional position. However, there's nothing inherent to a sports jacket that should mean professor or professional. Rather, it's only a system of convention, just like the way that the word "tree" only by convention means those leafy things we see outside. Similarly, my jacket only means a kind of professionalism insofar as it's not a t-shirt, or jeans, or a leather jacket, which I might wear under other circumstances to mean something else, like I'm going to a party."

Example 3: *Classroom*
"One way to look at this would be through the example of a classroom, which is in itself a kind of language, or system of meaning and communication. In a classroom, there is a relationship between the desks in which students sit, and the lectern in which the professor stands. There's nothing inherent to the lectern which should mean that it's where a professor goes, or the desk that means where to, where students sit, but rather it's only in their relationship that the desk gains meaning. It's in that relationship of difference in other words, that, um, the very positions of student and professor come to make sense. Through that system, or structure, of difference. This is a semiotic insight."

Example 4: *Chess Board*
"One way to look at this would be through the example of a chess board, which is also quite a lot like language. There is nothing inherent to the particular piece called a pawn that means it can only move one part on the board. It's only within that system, called chess, that the pawn has that kind of function. It's therefore the relationship between the ch– between the pawn and other figures: the rook, the queen, the king for example that we know what the meaning of the pawn is. It's only through those differences that any word gains meaning."

Example 5: *Traffic Lights*
"One way to look at this is through the example of a traffic light. So think for example of the colour red, which we know means "stop." However there's nothing inherent to the colour red that would mean "stop," it's only within that system of the traffic light that it would make sense. So, semiotics would argue that when we look at the red on the traffic light, we are actually saying "oh, it's not green and it's not yellow." It's only that relationship between the colours in a system through which red means "stop" through which the word "tree" gains meaning."

APPENDIX B.

VALID AND INVALID NOVEL EXAMPLES FOR TRANSFER TEST

Semiotics: Valid

"One way to think about semiotics is by thinking about it through nonverbal communication, like a thumbs-up. There is nothing inherent about a thumbs-up that means "good," however in North American culture, that is what it means. This does not mean that "thumbs-up" cannot mean something else in a different context, but its meaning is constructed through the situation that it is communication. A "thumbs-up" only gains meaning to the observer through its relationship with the culture that it exists within."

Semiotics: Invalid

"One way to think about semiotics is through the facial expression of sadness. It conveys a person's emotions and can be transmitted throughout different contexts. There is a biological correlate to sadness, which makes the facial expression partially inherent, while other parts of it may be learned. The facial expression of sadness transmits meaning to the observer through its relationship to other facial expressions that the individual may have."