

# Re-Categorization: Restructuring In Categorization

David G. Cosejo (dcosej1@uic.edu)  
Justin Oesterreich (joeste1@uic.edu)  
Stellan Ohlsson (stellan@uic.edu)

University of Illinois at Chicago  
Chicago, IL 60607 USA

## Abstract

The process of overriding prior experience and learning – *restructuring* – was examined within a categorization paradigm. Undergraduate students worked on a modified categorization task in which they learned an initial “misconception” category and later had to restructure their knowledge to acquire a different target category. Participants were able to learn both the misconception and the target category. The data are consistent with the conclusions that participants searched the relevant attribute space in parallel, that negative feedback had little effect, and that the supposedly rejected “misconception” was still affecting behavior at the end of target category training.

**Keywords:** restructuring; categorization; cognitive change.

## Introduction

The world we live in is constantly changing. Changes may be dramatic and large-scale, exemplified by the way Hurricane Katrina devastated New Orleans and the way fiscal crises change the global economy. Most changes are smaller and less destructive. We might have to adjust our commute due to changing train schedules, or we might have to learn to use a new computer. At both the global and the mundane scales, our environment is changing and we must change accordingly.

There are multiple lines of work in the cognitive sciences that address the issue of adaptation to changing task environments. In the area of skill acquisition, researchers model how people learn to perform unfamiliar tasks through practice (Ohlsson, 2008) and try to measure the mental cost of task switching (Altmann, 2007). In studies of creativity and insight, participants are faced with problems that require them to override their initial task representation (Kershaw & Ohlsson, 2004; Öllinger, Jones & Knoblich, 2006). Educational researchers investigate how to promote conceptual change in science learning (Ohlsson, 2009; Sinatra & Pintrich, 2003). Social psychologists have proposed mechanisms for attitude and belief change (Gawronski & Bodenhausen, 2006; Petty et al., 2006).

The central unsolved theoretical problem is how someone can learn something new that conflicts with what they knew before (Ohlsson, 2009). If new information is understood with the help of already acquired concepts, how, by what cognitive process, can a person acquire information that

contradicts those prior concepts? For brevity, we refer to this process as *restructuring*.

Why is restructuring difficult? In the science education context, one plausible hypothesis is that scientific concepts are more complex than common sense ones. The ordinary concept of “pulling” is simpler than the concept of “mutual attraction” that underpins mechanics; the concept of “mixture” is less complex than the chemical concept of “dynamic equilibrium”; and so on. Science concepts require attention to more aspects of a situation for correct application. Restructuring in science learning may be difficult because it requires the learner to move from a simpler to a more complex conception.

In this paper, we describe an experimental paradigm called *re-categorization* that allows us to bring restructuring into the laboratory for close-up study. We use a standard categorization paradigm, but change the target category once the participants start responding correctly. Adaptation to the change is stretched out over multiple categorization trials. Also, re-categorization affords experimental control over the relevant prior knowledge because participants learn their ‘misconception’ in the course of the experiment.

We describe the re-categorization paradigm in more detail, report an empirical study, highlight the salient features of participants’ behavior and discuss what cognitive mechanism may be able to reproduce and explain those features. Our results are consistent with a gradualist conception of the restructuring process.

## The Re-Categorization Paradigm

During a trial in the standard categorization paradigm, a participant sees a stimulus, predicts whether it is an instance of the category to-be-learned and receives feedback. Trials continue until the category is learned according to some mastery criterion (Ross et al., 2008). The re-categorization paradigm extends this task by incorporating the key feature of the Wisconsin Card Sorting Task (WCST), a clinical measure of perseverance: Once the category has been acquired, the experimenter shifts the learning target to another category *without telling the participant* (Stratta, et al., 1997; Obonsawin, 1999). Responses that were correct before the shift might now be incorrect and vice versa.

A re-categorization task thus consists of two phases: First, there is the initial training phase in which participants acquire the *initial category* (‘misconception’). This phase ends with the “behind-the-scenes” shift, and is followed by

a second training phase on the second, *target category*. The question is how and by what processes participants “back out” of the initial category and acquire the target category.

### Study

If participants solve categorization tasks by searching through the relevant hypothesis space, then a larger hypothesis space should increase the difficulty of the task. If a *simple category* is defined by a single value, a *feature*, on one of six binary attributes, then there are 12 possible hypotheses. But there are 132 possible *complex categories*, defined by a conjunction of two features. Hence, complex categories should be harder to learn. Also, category complexity should interact with restructuring. A simple-to-complex shift should be more difficult than either a simple-to-simple or a complex-to-complex shift. (It is less clear what to predict for a complex-to-simple shift.)



Figure 1: An example of a Martian bacterium.

**Participants** Participants in the present study were 278 undergraduates in an introductory psychology course who received course credit for participating.

**Materials** We created a set of 64 images of fictional micro-organisms ("Martian bacteria"). An example is displayed in Figure 1. The example has a tail with cilia, a single cell wall, a white cell body, one nucleus, ribosomes and three head bulbs. The possible attributes and their values are listed in Table 1. The materials also included a written instruction and a set of rating scales for recording participants' confidence judgments. All materials were presented via a computer screen and all participant responses were recorded using the E-Prime software package ([www.pstnet.com/products/E-Prime/default.htm/](http://www.pstnet.com/products/E-Prime/default.htm/)).

Table 1: Attributes and their possible values (features).

Attributes	Values (features)
Tail cilia	Present or not present
Number of nuclei	Two or zero
Ribosomes	Present or not present
Cell wall	Double or single layer
Head bulbs	Three or zero
Cytoplasm	White or grey

**Design** The study was a 2 X 2 design, with either a simple (one-feature) or a complex (two-feature) initial category and

either a simple or a complex target category. Participants were randomly assigned to the four conditions (labeled SS, SC, CS, CC), shown in Table 2. The specific features that defined the initial and target categories are displayed in Table 3.

Table 2: Design and number of participants in each cell.

Initial category	Target category
Simple	Simple
Simple	One→one (n = 73)
Complex	Two→one (n = 76)
	Complex
	One→two (n = 64)
	Two→two (n = 65)

**Procedure** Participants were asked to read instructions presented onscreen and to ask the experimenter if he or she had any questions. Participants read that the bacteria were discovered on Mars and that their task was to decide if they were “oxygen resistant” or “not oxygen resistant.” Participants saw a labeled variant (see Figure 2) and were instructed to “memorize the names of these six features.”

Table 3: Category definitions for the four conditions.

Condition	Initial category	Target category
Simple-to-complex (SC)	Head bulbs	Two nuclei & tail cilia
Simple-to-simple (SS)	Head bulbs	Two nuclei
Complex-to-complex (CC)	Ribosomes & single cell wall	Two nuclei & tail cilia
Complex-to-simple (CS)	Two nuclei & tail cilia	Head bulbs

Participants were asked to rate how important they thought the six attributes would be for predicting oxygen resistance. Each of the six attributes were rated on a scale from zero to five, with zero being “not important” and five being “very important.” Feature importance ratings were gathered three times: before initial training, after training on the initial category and again after training on the target category.

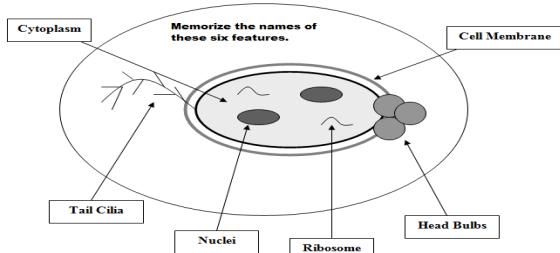


Figure 2: Labeled stimulus from the task instruction.

After reading the instructions, participants began training on the initial category. Training consisted of *training blocks* of 16 trials each. In each trial, participants saw a single image randomly selected from either the training subset or the assessment subset, depending on which type of block participants were in. The 64 images were split into 32 training images and 32 assessment images. Images in the assessment blocks were never included in training blocks. Participants then pressed either "Y" or "N" (yes or no) in response to the question "is this bacterium oxygen resistant?" and used the number keys to indicate their level of confidence in their decisions. Participants then received feedback (see below). Participants continued training until they reached a mastery criterion of 15 out of 16 correct responses in a training block. The maximum number of training blocks for the initial category was 10. Participants were then presented with an *assessment block* of 16 trials in which no feedback was given and all stimuli were previously unseen. At this point, participants were also asked to rate the importance of the six attributes.

When training resumed, the to-be learned category had changed. Participants were not told of this shift. The training procedure for the target category was the same as for the initial category. Participants were then presented with a second assessment block. They were again asked to rate the importance of the six attributes.

On each training trial, participants received feedback on whether or not their response was correct. Positive feedback following a correct response stated, "You are correct. This bacterium was indeed oxygen resistant." Negative feedback following an error stated, "You are incorrect, this bacterium was not oxygen resistant." Similarly for "not oxygen resistant." During the assessment blocks, participants did not receive feedback. At the end of the task, participants were debriefed and thanked for their participation.

## Measures

For each trial we recorded whether the response given was consistent with the initial category, whether it was consistent with the target category, a confidence judgment, response time and time spent looking at the subsequent feedback screen. We recorded the number of training blocks taken to reach criterion for the initial and target categories, respectively. We collected attribute importance ratings before training, after initial category training and after target category training.

We also computed a *change ratio* (CR) which examined the balance between the two types of responses. The CR is computed as the ratio of the difference between target and initial responses to their sum (within a block of 16 trials):

$$CR = (Target - Initial) / (Target + Initial)$$

The CR variable is zero when a participant's responses are equally consistent with both categories, -1 when his or her responses are completely consistent with the initial category

and +1 when completely consistent with the target category. Responses inconsistent with both categories do not affect the ratio.

## Results

**Acquisition of the initial category** The number of participants who acquired the initial category ('misconception') varied greatly across conditions. Of the 137 participants who were trained on a simple, single-feature initial category, 89 (65%) reached the mastery criterion in less than 10 training blocks. However, of the 141 participants who were trained on a complex, two-feature initial category, only 11 (8%) learned it. Thus, complexity mattered initially [ $\chi^2 (1, n = 100) = 60.84, p < .001$ ]; see Table 4. The two conditions with complex initial categories (CS and CC) did not produce sufficient participants who acquired the initial category for meaningful data analysis, so the following analyses focus on the two conditions with simple initial categories (SS and SC).

**Acquisition of the target category** Of the 89 participants who learned the initial category, 74 (83%) also successfully restructured their definition of "oxygen resistance" and reached the mastery criterion for the target category as well in less than the maximum of 10 training blocks; see Table 4.

Table 4: Number of participants who learned initial and target categories by condition. Percentages refer to the column immediately to the left.

Condition	n	Initial category (%)	Initial and target (%)
SC	64	41 (63.0)	39 (95.1)
SS	73	48 (66.0)	35 (72.9)
CC	65	7 (11.0)	6 (85.7)
CS	76	4 (5.0)	2 (50.0)
Collapsed	278	100 (36.0)	82 (82.0)

Did success rate differ, depending on the complexity of the target category? Of the 89 participants who learned the single-feature misconception, 35 were trained on the single-feature target category and 39 were trained on the two-feature target category. Contrary to our predictions, there was no effect of target complexity. For the simple target, 35 of 48 (73%) acquired the target category, while for the complex target category, 39 of 41 (95%) acquired the target,  $\chi^2 (1, n = 74) = 0.78, ns$ .

How much cognitive effort did restructuring require? Figure 3 shows the frequency distribution of the number of blocks to successful acquisition of the target category for all 89 participants who learned the initial category. Most of the participants had restructured by Block 3, (i.e., within 48 training trials). The modal value was Target- 2 or within 24 trials. However, the distribution is strongly skewed, with a long tail. Notice that the frequency shown for the 10th

training block includes the participants who had not yet learned the target category by the end of Block 10.

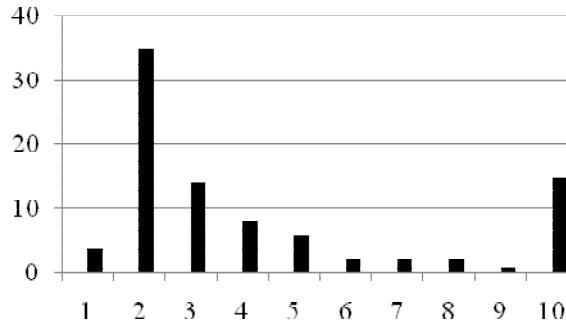


Figure 3: Frequency distribution for number of blocks to restructure for the SC and SS conditions (n = 89).

Did the amount of cognitive effort differ as a function of complexity? Once again, the answer is no. There is no significant difference between the mean number of blocks to restructure for the SC and SS conditions,  $t(74) = -0.36, ns$ .

To study the change from the initial to the target category in more detail, we plotted the CR measure (see the Measures section) separately for the SC and SS conditions; see Figures 4 and 5. In Figures 4 and 5, the mean CR score for each block *after* the shift is displayed. ‘Target-1’ indicates the first target training block after restructuring and so on. ‘A2’ indicates Assessment Block 2.

On the first assessment block (A1, not shown), the mean CR score is close to -1 because this analysis only includes participants who successfully acquired the initial category. Participants in both conditions are still responding below zero (i.e., in accordance with the initial category) in Target-1 after the shift in the to-be-learned category, but both are above zero by Target-2, indicating rapid adaptation to the change in feedback in both conditions.

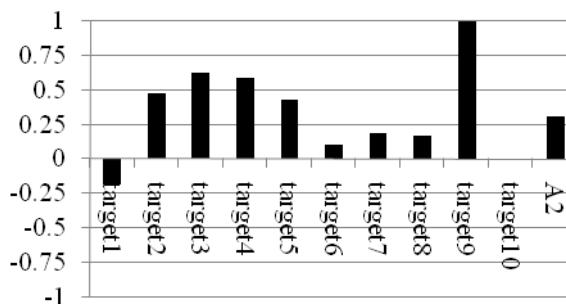


Figure 4: Mean CR for the SC condition by target training blocks (n = 39).

In Figure 5, the SS condition exhibited relatively steady, almost linear progress and all participants had restructured by Target-8. In the SC condition, progress was nonlinear, indicating that participants varied in how quickly they

restructured, and some participants had only completed restructuring by Target-9. It is plausible that SC participants had to realize that the target category was defined by two features, and they varied in how quickly they reached this insight. Both conditions scored a perfect +1 in the last training block because this analysis only includes successful participants.

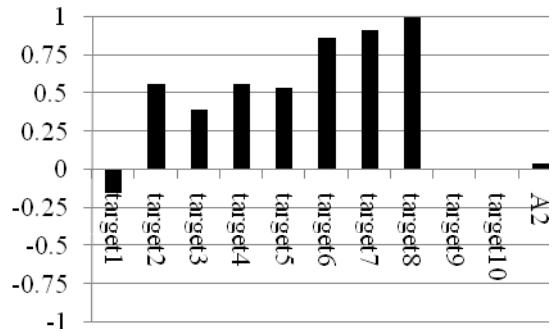


Figure 5: Mean CR for the SS condition by target training blocks (n = 35).

What did the participants learn? What was their state of knowledge once they had learned the target category? The obvious hypothesis is that they replaced the initial definition of oxygen resistance with the target definition in response to the negative feedback on erroneous postshift decisions. The data do not support this *replacement hypothesis*. First, performance on Assessment Block 2 (A2) is close to zero on the CR score in both conditions. That is, once feedback was removed and they were asked to categorize stimuli that were previously unseen, their performance became partially consistent with the initial category again (Figures 4 and 5; see the values for A2). Recall that a CR score of zero means that half the participant's responses were consistent with the initial category and half with the target category.

Second, the feature importance ratings do not support the replacement hypothesis. Figures 6 and 7 show the three sets of importance ratings broken down by the six attributes. Recall that ratings were collected before training, after initial training and after target training.

The middle set for both Figures 6 and 7 confirms the successful learning of the initial (“head bulbs”) category. Head bulbs are judged as more important than in the first set and all other features have decreased in importance. However, after *successful* learning of the one-feature target category “two nuclei,” the head bulb feature has decreased in importance again, but the correct feature is accorded *less* importance than it was initially, and the irrelevant “shaded cytoplasm” feature is considered almost as important (Figure 7). Notice that no feature is rated zero. This pattern is repeated for the two-feature target category (Figure 6). The “head bulbs” dominate the second set of importance ratings, but in the third set, the correct features, “two nuclei”

and “tail cilia”, are awarded less importance than they were initially, and the “head bulbs” feature is still considered somewhat important. Again, no feature was rated zero.

We examined whether feature importance ratings changed across the three occasions on which participants were asked to give ratings. Within-subjects ANOVAs with three levels of time revealed that in the complex condition, all six features changed in importance over time, “nuclei,”  $F(2, 39) = 23.38, p < .001$ , “ribosomes,”  $F(1.72, 39) = 69.46, p < .001$  with Greenhouse-Geisser correction, “cilia,”  $F(2, 39) = 26.31, p < .001$ , “head bulbs,”  $F(1.34, 39) = 39.67, p < .001$  with Greenhouse-Geisser correction, “cell wall,”  $F(2, 39) = 35.18, p < .001$  with Greenhouse-Geisser correction, “shaded cytoplasm,”  $F(2, 39) = 47.96, p < .001$ . Similar analyses were conducted and similar results were obtained for the simple condition. In short, participants’ feature importance ratings varied as a function of time as participants first learned the initial category and then the target category.

Furthermore, we asked whether there were differences in importance ratings among the features within each occasion. At each time, we treated each feature as a within-subjects factor. For the complex condition at Time 1, importance ratings were different from each other,  $F(3.81, 39) = 4.31, p < .01$  with Greenhouse-Geisser correction. Similar analyses were conducted and similar results were obtained for Time 1, 2 and 3 for both conditions.

## Discussion

The data suggest two conclusions with respect to the final knowledge state of the participants. First, their final conception of oxygen resistance is not well described as a conjunction of features. Instead, features appear to be evaluated individually, with participants assuming that more than one feature may be relevant. This is why the two-feature target category was not radically more difficult to acquire. Participants considered all features in parallel in both conditions. Second, the features were not evaluated on an either-or basis for relevance, but are better described as having a degree of importance. Third, there was no complete rejection of any feature. Both the initial category definition and the features that were irrelevant throughout the task were awarded some degree of importance at the end.

What does this characterization of the final knowledge state imply about the mechanisms of change? What type of learning mechanisms would have this kind of outcome? A first-approximation explanation follows, if we assume that the task is awarded a fixed amount of a quantity that we may call ‘strength.’ Strength gets assigned to the different features depending on the stimuli seen and the feedback given. As participants go through initial training, almost all the strength is allocated to the “head bulbs” feature. To explain what happens next, we hypothesize that negative feedback was ineffective. The allocation of strength across

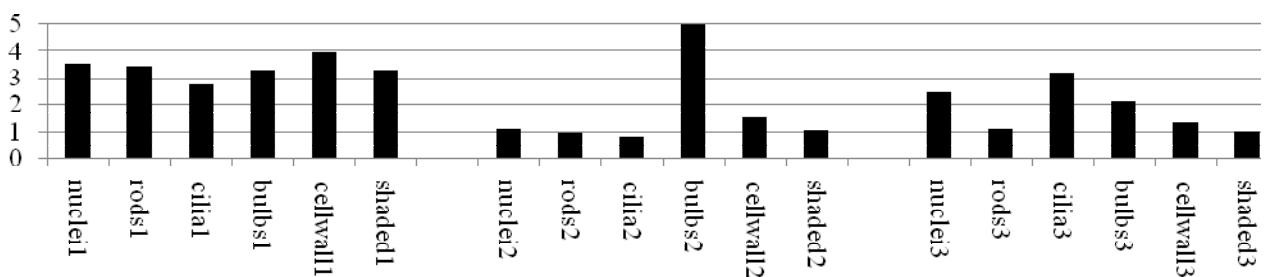


Figure 6: Feature importance ratings for SC condition (n = 39).

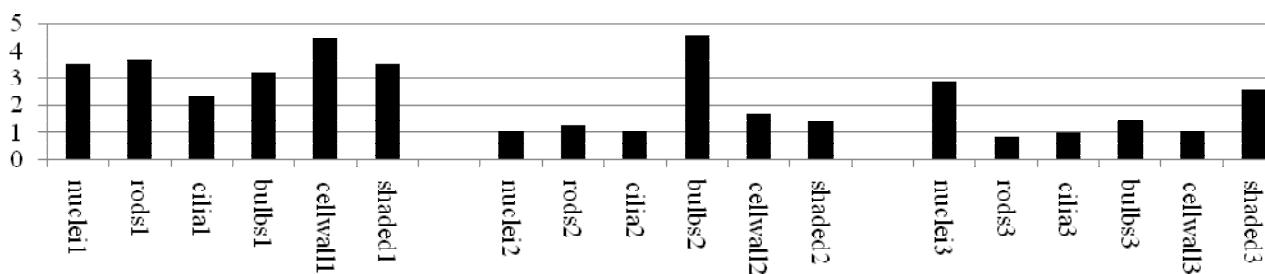


Figure 7: Feature importance ratings for SS condition (n = 35).

the features is only adjusted when there is positive feedback. However, the increase of strength of one feature affects that of other features because the total amount of strength is fixed. Thus, as learning of the target progresses, there is a gradual increase in the strength of the relevant features while the other features lose some of theirs. However, to be consistent with the pattern of importance ratings, we have to further assume that the loss of strength is proportional to how much strength a feature already has. Most of the loss occurs on the “head bulbs” feature and less on the other features. The overall effect is that the strength is spread out across the features, instead of concentrated on the correct definition of the target category. The target features do not ever reach the level of the “head bulbs” feature because there is not enough strength to go around. The result is a final knowledge state in which all the features have some strength and the differences between the target features and the irrelevant features are relatively small. This explains the inconsistent performance on the second assessment block.

An account along these lines contrasts with multiple expectations derived from the idea that people learn abstract, rule-like representations of the relevant categories in this type of experiment. First, it is plausible to assume that people treat conjunctive categories as conjunctions of select features, while the pattern in our data is more consistent with the idea that participants treat all the features as relevant to varying degrees. Second, it is plausible that people restructure in response to contradictory feedback, while our data are consistent with the idea that only positive feedback has an effect. Third, it is plausible that after restructuring, the old knowledge structure has been changed or deleted, but our data indicate that it is still active; interestingly, researchers in other fields have reached analogous conclusions (Ohlsson, 2009; Petty et al., 2006).

The present study had several shortcomings and limitations that will be addressed in future work. The informal explanation above needs to be implemented as a simulation model in order to investigate whether it can reproduce the quantitative details of the data. The counterintuitive lack of a complexity effect needs to be replicated with other stimuli and other behavioral data such as think-aloud protocols to see whether it is robust and generalizable. In particular, the effect needs to be replicated with realistic stimuli and categories. Finally, the exact implications for areas like insight, conceptual change and belief revision remain to be worked out.

### Acknowledgments

Preparation of this paper was supported by Award N000140910125 from the Office of Naval Research. We thank Cyndie Penrose, Tim Sparer, Amana Mohammad, and Noah Cepela for assistance with data collection and analysis. Further thanks go to Bettina Chow for helping pioneer the recategorization paradigm.

### References

Altmann, E. M. (2008). Comparing switch costs: Alternating runs and explicit cueing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 475-483.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692-731.

Kershaw, T. C. & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(1), 3-13.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1534-1555.

Obonsawin, M. C., Crawford, J. R., Page, J., Chalmers, P., Low, G., & Marsh, P. (1999). Performance on the Modified Card Sorting Test by normal, healthy individuals: Relationship to general intellectual ability and demographic variables. *British Journal of Clinical Psychology*, 38(1), 27-41.

Ohlsson, S. (2008). Computational models of skill acquisition. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (pp. 359-395). Cambridge, UK: Cambridge University Press.

Ohlsson, S. (2009). Re-subsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist*, 44, 20-40.

Petty, R., Tormala, Z., Brinol, P., & Jarvis, W. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality and Social Psychology*, 90, 21-31.

Ross, B., Taylor, E., Middleton, E., & Nokes, T. (2008). Concept and category learning in humans. In H. L. Roediger, III (Ed.), *Cognitive psychology of memory*. Oxford, UK: Elsevier.

Sinatra, G. M. & Pintrich, P. R. (2003). *Intentional Conceptual Change*. Philadelphia: Lawrence Erlbaum.

Stratta, P., Daneluzzo, E., Prosperini, P., Bustini, M., Mattei, P., & Rossi, A. (1997). Is Wisconsin Card Sorting Test performance related to working memory capacity? *Schizophrenia Research*, 27(1), 11-19.

Öllinger, M., Jones, G., & Knoblich, G. (2006). Heuristics and representational change in two-move matchstick arithmetic tasks. *Advances in Cognitive Psychology*, 2, 239-253.