

# Computational Models for Representation Change in Human Learning

Jennifer Roberts

CSAIL, Massachusetts Institute of Technology; jenmarie@mit.edu

To develop computational models for human learning and development, we must understand how humans organize information about their environment and experiences. Intuitively, representation change plays a key part in this process, because changes in representation allow us to model how the world works, organize information into comprehensible units, and identify pertinent information so that we can ignore irrelevant details. To develop models of human learning that account for representation change, we must first identify classes of representation change commonly encountered during learning and development, analyze how mechanisms for representation change might interact with other mental processes, and identify cognitive constraints that might lead to representation change. Based on this analysis, we can then move toward identifying key components that a computational model for human representation change would require and analyzing the validity of current approaches within this context.

In this paper, I identify representation changes observed during human learning and development, suggest a working model of how representation change mechanisms might interact with other cognitive subsystems, and suggest computational mechanisms that might account for the types of representation change observed in humans.

## Human Evidence

Table 1 presents several types of representation change that collectively address aspects of word learning and abstract concept creation (Niyogi, 2005), chunking strategies for data compression (Ericsson, Chase, & Faloon, 1980), theory formation (Tenenbaum, 2000), and analogical reasoning (Gentner & Markman, 1997).

## Working Model

Figure 1 shows a proposed working model that describes how representation change mechanisms might relate to other cognitive mechanisms. During passive perception, a person's sensory mechanisms receive large amounts of data from the outside world, much of which does not receive attention. During active information processing, the person attends to some portion of the new information and incorporates noteworthy information into existing representations.

Representations that capture patterns, or theories, play an integral part in inference and prediction. For example, once a child decides that the seasons follow a cyclic pattern, she will infer that fall always follows summer.

Types of Representation Change	Examples
Word set to Map from words to sensory percepts or concepts	<ul style="list-style-type: none"> <li>- List of color words to map between word and color percept</li> <li>- Sweet and sour to 4-year-old's observation that "sour makes your taste buds stand up"</li> </ul>
Unordered set to Ordered set; Creation of ordered sets	<ul style="list-style-type: none"> <li>- Numbers, letters</li> <li>- Relative size, height</li> <li>- Casual theories</li> </ul>
Ordered/patterned set to Spatial construct	<ul style="list-style-type: none"> <li>- Mapping time to a line</li> <li>- Colors to a circle</li> <li>- Seasons and days of the week to a cycle</li> <li>- Elements to a periodic table</li> </ul>
Unordered set to Groups of similar concepts Chunking strategies	<ul style="list-style-type: none"> <li>- Mnemonic devices (acronyms, songs, associations with known concepts)</li> <li>- Outlining topics and subtopics</li> <li>- Grouping based on features (shape, function, etc.)</li> <li>- Grouping based on co-occurrence</li> </ul>
New relationship to Association with similarly structured relationship	<ul style="list-style-type: none"> <li>- 4-year-old's observation that "Baltimore is in the tummy of Maryland"</li> <li>- Electrons move around the nucleus like the earth moves around the sun</li> </ul>
Single concept to Separate concepts with greater explanatory power	<ul style="list-style-type: none"> <li>- Sleep and death as a single concept to sleep and death as separate concepts</li> </ul>

Table 1 – Human Evidence of Representation Change

Based on internal theories, people continually make inferences about the world around them. When these inferences produce inaccurate results, they often attract a person’s attention. When this occurs, people often start storing observations that are inconsistent with their theory until they amass enough inconsistent information to warrant developing a new theory.

As people incorporate information into existing representations, eventually the representations will become cumbersome to work with, leading to a sense of cognitive overload. This overload may be due to (1) working memory constraints, i.e., the inability to hold all relevant information in working memory while performing a task, (2) long term memory constraints, i.e., the inability to store or recall poorly-organized information, (3) repeated observation of important-to-understand information that cannot be explained by current theories, (4) redundant observations, and (5) detection of inconsistencies within current theories. To reduce the cognitive overload, the person develops a new representation. This process might include (1) recognizing patterns and similarities, (2) clustering information into small sets, (3) creating new constructs, and (4) mapping new information to existing constructs. A representation change might then prompt someone to test the new representation, which would generate more perceptual data.

### Computational Mechanisms for Representation Change

Based on the human evidence presented in Table 1, a model of representation change should include computational mechanisms that organize memory into sequences and chunks, create associations, create new theories, and update existing theories. Table 2 suggests computational mechanisms that might be involved in these processes. Because these mechanisms perform a diverse set of operations, we need multiple learning frameworks and algorithms to model human learning and development.

Representation Change	Computational Mechanisms
Memory Organization	<ul style="list-style-type: none"> <li>- Map new info to existing info (mnemonic devices)</li> <li>- Posit new nodes to chunk data</li> <li>- Sequentially order data</li> </ul>
Development of Associations	<ul style="list-style-type: none"> <li>- Detect correlations</li> <li>- Create association between correlated words, percepts, experiences, or abstractions</li> </ul>
New Theory Creation	<ul style="list-style-type: none"> <li>- Detect patterns (cycles, casual relationships)</li> <li>- Use patterns to compress data</li> <li>- Use compressed representation to perform inference and prediction</li> </ul>
Theory Modification	<ul style="list-style-type: none"> <li>- Perform inference using current theory</li> <li>- Use observations that agree with inferred results to strengthen confidence in theory</li> <li>- Identify and store observations that conflict with theory</li> <li>- Detect significant evidence against theory; use to develop new theory</li> </ul>

Table 2 – Computational Mechanisms for Representation Change

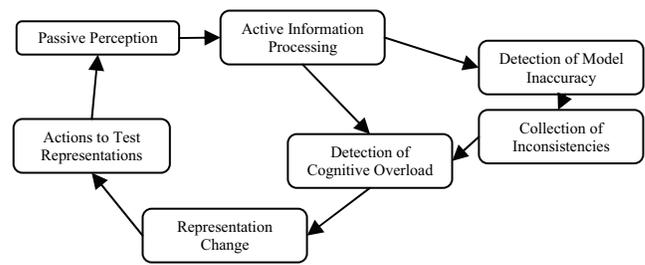


Figure 1 – Working Model of Representation Change

### Contributions

In this paper, I presented an initial list of representation changes observed in humans, proposed mechanisms for triggering representation change, and illustrated the need for a diverse set of computational learning mechanisms, thus specifying a set of minimum requirements necessary for a comprehensive computational model of representation change in human learning and development.

### Acknowledgements

I would like to thank the Fannie and John Hertz Foundation for their financial support.

### References

Ericsson, K.; Chase, W.; Faloon, S. 1980. Acquisition of a Memory Skill. *Science* 208(4448): 1181-1182.

Gentner, D.; Markman, A. 1997. Structure Mapping in Analogy and Similarity. *American Psych.* 52(1): 45-56.

Niyogi, S. 2005. Steps Toward Deep Lexical Acquisition. In Proc. of the 2<sup>nd</sup> Workshop on Psychocomputational Models of Human Language Acquisition, 91-99. Ann Arbor, Mich: Association for Computational Linguistics.

Tenenbaum, J. B. 2000. Rules and Similarity in Concept Learning. *Advances in Neural Information Processing Systems* 12 59-65. Cambridge, Mass: MIT Press.