Title
Coherence in the Visual Imagination: Local Hill Search Outperforms Thagard's Connectionist Model

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 36(36)

ISSN
1069-7977

Authors
Vertolli, Michael
Davies, Jim

Publication Date
2014

Peer reviewed

Coherence in the Visual Imagination: Local Hill Search Outperforms Thagard's Connectionist Model

Michael O. Vertolli (michaelvertolli@gmail.com)
Jim Davies (jim@jimdavies.org)
Carleton University, Institute of Cognitive Science, 1125 Colonel By Dr.
Ottawa, ON K1S 5B6 Canada

Abstract

A cognitive model of the visual imagination will produce "incoherent" results when it adds elements to an imagined scene that come from different contexts (e.g., "computer" and "cheese" with "mouse"). We approach this problem with a model that infers coherence relations from co-occurrence probabilities of labels in images. We show that this algorithm's serial traversal of networks of co-occurrence relations for a particular query produces greater coherence than one leading model in the field of computational coherence: Thagard's connectionist model.

Keywords: imagination; coherence; artificial intelligence

Introduction

The imagination is implicated in a wide range of abilities related to human cognition. The list of abilities includes but is not limited to planning, problem solving, hypothetical thinking, counterfactual thinking, theory of mind, and mental time travel (Davies, Atance, & Martin Ordas, 2011). Despite a plethora of research on the imagination as a facilitator for these abilities (see, for example, Markman, Klein & Suhr, 2012), the generative capacity of the imagination is an untapped area of research in this domain. This work focuses on the visual faculty as it is the most studied.

When someone constructs a visual scene with the imagination (e.g., a mouse eating a piece of cheese) they might use visual memories from many different experiences as the components of the new scene. How these components are selected from the range of possible experiences is not obvious. If more than just the mouse and cheese are included in the scene, it is unclear what makes the selection of some elements (e.g., a cat, mousetrap, floorboards, or a countertop) more likely or appropriate than others (e.g., a rollercoaster, map of Spain, or cruise ship).

What is known is that people do not *arbitrarily* select the components for their imaginings, even if those imaginings are entirely fictional (see Cockbain, Vertolli, Davies, in press). There is an intuitive coherence imposed on imagined scenes that inhibits unusual and sometimes even highly creative combinations.

One way that humans might make this selection is through the co-occurrence of objects in visual memory (by visual memory we mean only the memory of visual things, and not a specific subsystem like the visuo-spatial sketchpad). Thus, when one is imagining a scene given an environmental query (e.g., a novel, question, or problem), mental processes might search visual memory for other objects that often occur with that query.

Recent research in cognitive neuroscience supports this co-occurrence view. Under the title of "scene construction theory," this research has demonstrated that the hippocampus plays a primary role in the construction of a "coherent spatial context" for the integration of the components of imagined experiences among other cognitive phenomena (Hassabis & Maguire, 2007; Maguire & Mullally, 2013). Parallel research on the role of the hippocampus in the emergence of conceptual knowledge and knowledge transfer (see Kumaran, Summerfield, Hassabis & Maguire, 2009) endorses the "memory space" hypothesis of hippocampal function. In this view, neurons in the hippocampus encode the co-occurrence of the components of a given experience or event through their spatiotemporal associations (Konkel & Cohen, 2009). This means that objects can co-occur as a consequence of associations due to spatial relationships (e.g., mice often are seen next to cheese) or temporal relationships (e.g., gunshots are often followed by death or injury). We only address the former in the current work.

Figure 1. An incoherent scene generated by SOILIE for the query 'mouse,' containing elements from both the computer and animal senses of 'mouse.'

In sum, when one imagines a mouse eating a piece of cheese, it is not surprising that mousetraps, cats, etc. are more likely to come to mind than unrelated elements. They all often occur together in the world and, thus, are associated in the brain. In this case, "a mouse eating a piece of cheese" serves as the query and the other elements are what are returned by some imagination process. One way to further explore this idea is through the use of computational models.

The Science of Imagination Laboratory Imagination Engine (SOILIE) is a computational model of a functional description of processes in the imagination that generate scenes from an environmental query (Vertolli, Breault, Ouellet, Somers, Gagné, Davies, 2014). In place of human 'experiences' and 'objects' SOILIE uses labeled images from the web. When generating a novel scene, SOILIE must determine which labels are appropriate to select, given a particular query. And, in keeping with the descriptions given above, SOILIE currently uses co-occurrence relations to make this selection. In this context, co-occurrence is determined by the frequency with which one label is present in the same image with another label.

SOILIE derives these co-occurrence relations from the Peekaboom database of labelled images. With over fifty thousand images and ten thousand labels, the Peekaboom database is one of the largest of its kind. The dataset is the combined result of two online games: the ESP Game and Peekaboom (Von Ahn, Liu, & Blum, 2006). In the ESP Game, pairs of players are shown the same image and without communicating try to enter the same words (Von Ahn & Dabbish, 2004). Words that both players enter are associated with the image and, consequently, common labels are applied to images collected from the internet. To prevent a narrow set of the most common words, labels would become unusable after repeated use. This increased the diversity and size of the resulting label set for each image.

SOILIE's dataset comes from a related game, Peekaboom, which uses ESP game data and results in images with labelled selections of pixels. Both games are designed to produce data that can be used in vision research. Thus, they are particularly relevant for SOILIE's task.

SOILIE uses co-occurrence probabilities extracted from the Peekaboom database described above. Co-occurrence probabilities are calculated by dividing the total number of images (*I*) in the Peekaboom database that contain the co-occurring label (*l*) *and* a particular query (*q*) by the total number of images with *just* the query. Using set theory notation, this yields:

$$P(l \mid q) = \frac{|I_q \cap I_l|}{|I_q|}$$

where ∩ indicates set intersection and || indicates cardinality (i.e., the total number of elements in the set). One important feature of this formalization is that it is non-commutative (i.e., *P* yields a different value for mouse-cheese than it does for cheese-mouse). Parallel research on co-occurrence in the machine learning literature suggests that this is both more realistic (e.g., most weddings have flowers but most flowers are not in weddings) but most models do not account for it (see Huang, Yu & Zhou, 2012; Zhang & Zhou, 2013).

Research in neuroscience suggests that visual working memory can hold approximately three to five objects of average complexity (i.e., they leave open the possibility that very simple objects and very complex objects might be affected differently; Cowan, 2001; Edin, *et al*., 2009). Thus, it is assumed that on average an imagined scene has approximately three to five elements in it at any given time; though aggregates (i.e., combining two or more elements into a single element) are entirely possible. Similarly, four labels, excluding the query, are retrieved by SOILIE from the co-occurrence data and five labels in total are selected for every imagined scene. We decided that this number, despite being in the upper part of the range, was the most useful: preliminary research suggested that larger sets of labels increased the divergence in the success of the underlying subsystems, five is still in the accepted range, and the query does not really need to be maintained in working memory to the same degree (an individual could always re-query) nor does it need to be retrieved.

However, after working with earlier instantiations of SOILIE (see Breault, Ouellet, Somers, & Davies, 2013), a problem became apparent. When images are selected purely on the basis of their co-occurrence with the initial query, that is, selecting the labels with the highest co-occurrence or "top-*n*," the scenes produced are often contextually incoherent in an intuitive sense.

For example, SOILIE was queried with the word 'mouse,' which is polysemous (i.e., it has multiple, related meanings; e.g., a computer mouse and the animal mouse). Each meaning of a polysemous word is represented by different images in the database—assuming that a single image would be highly unlikely to contain both meanings of a given label (e.g., have the animal on a desk with a computer). Each of these different images is similarly associated with a different collection of labels and each set of labels has a different set of co-occurrence relations. Problematically, by being reduced to a collection of co-occurrence probabilities in visual memory, the sets of images, labels, and co-occurrence relations that separate the two polysemous meanings of the word "mouse" are no longer directly detectable. They are collapsed into a single dimension associating pairs of labels (see Table 1).

The result is that models, like SOILIE's original "top-*n*" model, which act on this single dimension, will be unable to infer the appropriate contextual distinctions from the differences in the underlying images. Thus, they will often produce incoherent images (see Figure 1).

The problem of 'coherence' is not exclusive to SOILIE. Generally, models that address context need to be able to select coherent combinations (Hullett & Mateas, 2009). Most models, due to memory limitations, will also need to reduce the original input (e.g., images) into some form of compressed data (e.g., co-occurrence probabilities). In this compression, some of the information will be lost. Thus, there will almost always be the dual task of 1) finding compression techniques that are better able to capture greater quantities of salient information and 2) build decompression procedures that can re-derive information that was lost through higher-order patterns of the remaining data.

Table 1: Label co-occurrence probability of two images alone and in SOILIE's complete database

Image 1 labels:
mouse, eye, rodent, rat, animal, ear, ears

Image 2 labels:
mouse, wires, monitor, screen, headphones, computer

Co-occurrence of each label with query "mouse" given only those two images: 0.5

Co-occurrence of label with query "mouse" using all images in the database:

rat	0.29	rodent	0.08
ear	0.19	ears	0.07
computer	0.17	eye	0.06
animal	0.13	headphones	0.01
monitor	0.12	wires	0.01
screen	0.10		

Top-4 labels from the two images using co-occurrence from database:
rat, ear, computer, animal

SOILIE's coherence subsystem, Coherencer (Vertolli & Davies, 2013), falls into the second category. Thus, it is not a memory system, *per se*, but rather defines a form of contextual optimization of a given type of memory system, mainly, one that uses co-occurrence as per SOILIE and the hippocampus. This characterization is consistent with both the cognitive neuroscience literature on the hippocampus, which sees the hippocampus as a system of constructive integration of memory components and with Paul Thagard's (2000) research on coherence.

Thagard (2000) describes the problem of coherence as an optimization problem. That is, given a particular structure (e.g., co-occurrence), coherence is the dynamic construction of the best combination of components to maximize or minimize a particular set of criteria. Thagard takes these criteria to be a set of positive constraints (i.e., inclusion of one component increases the likelihood of inclusion of another component) and negative constraints (i.e., inclusion of one component decreases the likelihood of inclusion of another component). These constraints are optimized by maximizing the number of positive constraints in a collection and minimizing the negative constraints.

After formalizing coherence in this way, Thagard proceeds to outline a number of general classes of computational models that can resolve this type of problem. He then dismisses all of them in favour of one: a connectionist algorithm. His argument, roughly, is that the parallel approach inherent in these algorithms is both better at finding the global optimum (i.e., the *best* coherence for a given set of constraints) and, for those algorithms that are comparable, it is more cognitively plausible. Thagard has implemented a number of such networks in related domains with success (e.g., Thagard, 1989, 1992, 2000; Eliasmith & Thagard, 1997). However, before continuing his exploration of his preferred algorithm, Thagard makes one caveat; mainly, that incremental algorithms, or what are commonly referred to as local hill searchers in the machine learning and optimization literature, might offer valuable insights into human cognition as both are known to perform sub-optimally in many domains, including coherence.

In what follows, SOILIE's Coherencer system will be compared to one instantiation of Thagard's connectionist algorithm in the current domain (i.e., visual coherence in the human imagination). In previous research, Coherencer was shown to better capture coherence than SOILIE's original, top-*n* model and a random search (Vertolli & Davies, 2013). It was also shown that Coherencer falls under Thagard's incremental class of algorithms. Thus, this comparison provides both a more robust test of Coherencer's efficacy, and insight into the cognitive implications that Thagard pointed to in his caveat.

It is worth noting that both Coherencer and Thagard's model, in as much as they exist in the brain, are *both* necessarily instantiated in neural processes. One should not confuse the semantic convenience of calling connectionist models "neural networks" with a literal network of neurons and Coherencer with "something else." What is being tested, then, is whether the higher-order functionality of hippocampal or similar processes is better replicated with a *serial* process or a *parallel* process, with the corresponding implications for optimality in the system (in as much as those implications are in fact accurate). A serial virtual machine can be implemented on a parallel computational architecture, neurological or otherwise. Thus, both types of processes are cognitively plausible when considering only this aspect.

Implementation

We will proceed by giving a very brief description of Coherencer (for a more detailed description, see Vertolli & Davies, 2013) and a detailed description of the current implementation of Thagard's connectionist algorithm.

Coherencer operates as follows. First, Coherencer creates a pool of all labels that co-occur with the query. From this pool, it initially puts the top-4 labels with the highest co-occurrence in its memory buffer. Then, a square matrix of all the labels is created where each cell holds the co-occurrence probability for the row-column pair or *P*(row(n), column(m)). The average of the entire matrix is calculated and, if it passes a threshold (λ)[1], the collection is accepted (i.e., if $\frac{1}{20}\sum_{n=1}^{5}\sum_{m=1}^{5} P(l_m \mid l_n) > \lambda$). We ignore the diagonals, where $n = m$, in this calculation. If it fails to pass the threshold, the label with the lowest co-occurrence with all other labels (i.e., for the *i*th label, the sum of all row(i)

[1] We take the threshold to be a coarse representation of a learned sense of coherence in the world. This means that the number would vary depending on one's experiences (i.e., given a different database) or in different contexts (e.g., when trying to be creative).

values and column(i) values) is discarded and cannot be reselected. A new label is then swapped in from the pool and the process repeats until either the pool is empty or a set passing the threshold is found. If the pool is empty, Coherencer fails to create an imagined scene.

The construction of the connectionist model will proceed as described by Thagard (2000). A node is constructed for the query and every label co-occurring with the query. For every positive constraint between two labels, an excitatory link is constructed between the corresponding nodes with a weight equal to the co-occurrence probability. For every negative constraint (i.e., when two labels have a co-occurrence of 0), an inhibitory connection is constructed between corresponding nodes with a weight set to the average co-occurrence of all non-zero values (φ).[2] An initial activation (0.01) is assigned to each node with a special locked activation (1.0) on the query node. All nodes then have their activation updated in parallel[3] using the following formula[4]:

$$\vec{a}_{t+1} = \vec{a}_t(1-d) + f(\overrightarrow{net})$$

where $\vec{a}$ is a vector of all the node activations at time *t* with each label as a cell in that vector, *d* is a scalar decay parameter (0.05) that decrements each node at every cycle. The vector $\overrightarrow{net}$ is computed by matrix multiplication as per:

$$\overrightarrow{net} = \vec{a}_t W$$

where *W* is the weight matrix for the network with its rows corresponding to the node being updated and the columns corresponding to the linked nodes (i.e., neighbours of node *i*). The values at $W_{i,i}$ (i.e., the diagonal of the matrix) are set to 0 so the activation passed from a node to itself is 0. *W* also corresponds to Coherencer's co-occurrence matrix with all co-occurrence values of 0 set to φ. Finally, *f* from the original equation is a function that performs element-wise multiplication with a different number depending on the elements direction from zero as per this equation:

$$f(\overrightarrow{net}) = net_i x \begin{cases} x = a_{max} - a_i \text{ if } net_i > 0 \\ x = a_i - a_{min} \text{ if } net_i \leq 0 \end{cases}$$

where *x* is the variable multiplier, a_i is the *i*th value of $\vec{a}$, a_{max} is the maximum activation of a node (1.0), and a_{min} is the minimum activation (-1.0). After the update, each node is reduced to the maximum and minimum activation values if it exceeds them.

In the larger process, activations update until the average change in the sum of all differences is less than a threshold (θ) or until 500 iterations occur. The following equation illustrates the former:

$$\Delta\vec{a}_t = \frac{1}{10n}\sum_{t-10}^{t}\sum_{i=0}^{n}\left(\left|a_{t,i} - a_{t-1,i}\right|\right) < \theta$$

where Δa is the change in activation over the past 10 iterations, $a_{t,i}$ means activation at time *t* and node *i*, || here indicates absolute value, and θ is the threshold. The 4 labels with the highest activation are selected providing a sort of top-4 filter.

In what follows, we will describe an outline of the comparison of these two models.

The Comparison

The comparison followed the basic structure of a memory task, where a participant is given a collection of data; this data is compressed in memory, and then it is recalled. In our abstraction of this structure, each model is given a collection of images, which are compressed into co-occurrence probabilities in memory. The models are then tasked with recalling this information.

Unlike a memory task, the goal is not to test the bounds of human or animal functionality. Instead, it is to assess the efficacy of the decompression step that re-generates the coherence information that was lost during compression in memory. Thus, the quantity of images remembered is not what is of interest. It is a certain quality in the generated images, namely coherence. This quality can be tested quantitatively by determining if the elements (in this case, labels) selected by the model when given a particular label or query do in fact occur in one of the original images. If they do, then the original coherence information has been successfully re-generated from the compressed data.

It is worth noting that, outside of the methodological advantages just outlined, this conceptual method is also theoretically more plausible than approaches that do not account for memory. The research in cognitive neuroscience previously outlined suggests that the imagination, spatial navigation, and memory are all associated through the underlying functionality of the hippocampus. Thus, by testing the models through this sort of generative recall (an integrated imagination-memory process), one might better approach the mechanism that underlies all of these processes: cognitive generation proper.

In either case, we hypothesize that Coherencer will outperform Thagard's model in the current comparison. We anticipate that serial processes better capture the contextual transitions necessary to appropriately frame a given scene. And, the advantages of using a parallel, non-linear optimization process are lost when dealing with a single feature.

[2] This was found to be 0.14878295850321488. The rounding occurred where it naturally does in the Python computer language (double precision float). As a consequence, different languages may get slightly different results unless this is controlled.

[3] The nodes are updated serially but the results of those updates are not used until the next serial update of all nodes. Thus, the end product is a parallel process implemented on a serial machine.

[4] All formulas are vectorized implementations of those described by Thagard (2000). I chose to use row vectors instead of column vectors as this more closely mirrors Coherencer's implementation.

Method

There are two models that were compared: Coherencer and Thagard's model. The entire Peekaboom database was initially filtered to remove all images with fewer than five labels and any labels that only occurred on those images. A total of 8,372 labels and 23,115 images remained after this filtration. All of the remaining images were compressed to their corresponding co-occurrence probabilities.

Each of the 8,372 labels was run through both of the algorithms 100 times and the results were averaged. Each query plus four returned labels are the elements of a new generated scene. The results for each of the algorithms were assessed with regard to the original images. If at least one image in the test set contained the five labels that were selected by a particular algorithm, including the query, the algorithm scored one point. If there were no images containing the five labels, they did not score a point. The results on each of the labels were paired for comparison. The total number of points scored by a model where the other model failed to score a point (i.e., excluding labels where both models failed or both models succeeded) were used to compare Coherencer to Thagard's algorithm.

Results

As hypothesized, Coherencer had more successful matches than the connectionist algorithm. The statistical details are as follows.

McNemar's repeated measures chi-square test demonstrates that Coherencer performed significantly better than Thagard's algorithm, $\chi^2(1, N=8372) = 7.80$, $p = .006$, $\varphi = 0.44$. The average scores in each of the categories are listed in Table 2. In this test, model runs where Coherencer and Thagard's algorithm both fail or both succeed on a given query (i.e., the models perform identically) are ignored; thus, the comparison occurs between the runs where one model failed and the other succeeded and vice versa. All values are reported for completion and evaluation purposes. As is standard with chi-square tests, both the actual number of runs and the statistically expected number of runs for a given category are reported. Figure 2 shows the standard deviation for the 100 model runs for the values that are used in the comparison (i.e., when the models are *not* performing identically). Even when the difference between the two categories is the smallest, the result is still statistically significant, $p = .031$.

Discussion

The results support the idea that Coherencer generates elements that create a more coherent scene than Thagard's model. However, the intent is not to falsify Thagard's claim to the formal optimality of connectionist algorithms over incremental algorithms in the domains he considers, which are largely about higher-order epistemological relations and constraints. Co-occurrence probabilities are part of a much lower system. The fact that Thagard's theory could anticipate both categories of systems we believe lends credence to it.

The purpose of this comparison is to extend the theory in order to better comprehend the subtle nuances implicated within it. For example, under what conditions are incremental algorithms present? Here, the evidence suggests that low level (i.e., co-occurrence), low dimensionality (i.e., just co-occurrence probabilities), with high combinatoric load (approximately 3.42×10^{17} possible 5-label combinations) requires incremental, heuristic approaches. Assuming the connectionist, parallel approach is optimal, how might that incremental approach switch into a functionally parallel one? Or, does it only approximate a parallel approach, which forces sub-optimal solutions in higher-order domains? Thagard (2000) explicitly mentions the tendency for humans to make sub-optimal decisions and the potential association between incremental approaches and bounded rationality (Simon, 1991). This project supports this association.

Table 2: McNemar χ^2 calculation between Coherencer and Thagard's model.

		Coherencer failure	Coherencer success	Total
Thagard's algorithm failure	Actual	2099.0	1166.0	3265.0
	Expected	1222.2	2042.8	
Thagard's algorithm success	Actual	1035.0	4072.0	5107.0
	Expected	1911.8	3195.2	
Total	Count	3134.0	5238.0	8372.0

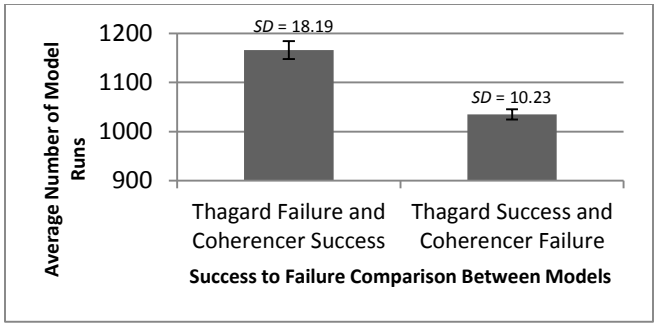

Figure 2: Failure-Success and Success-Failure average scores with standard deviation bars for 100 model runs

Research in working memory has also described a limited, serial system—the episodic buffer—that roughly matches what Thagard is describing (Baddeley, 2000). The episodic buffer is believed to be the means of integration for the different sense modalities as well as the retrieval mechanism for long-term memories. That is, it is mapped to a roughly identical, functional domain as the hippocampus. The limitations of these systems might result in downstream

limitations, and this suggests a rather simple explanation for bounded rationality.

In the current research, these observations suggest interesting implications for Coherencer. With respect to human cognition, Coherencer might better model the bounds of human rationality than the alternatives, including Thagard's connectionist models. Local hill searchers (i.e., incremental algorithms) might be optimal if the compression in memory reduces the feature space to a low dimensionality where non-linear techniques like Thagard's model give too little advantage for their increased cost in time and resources. The parallels with both the hippocampus and the episodic buffer suggest that Coherencer might also provide a functional model and computational implementation that better describes contemporary research in these domains than the competitors. Additionally, it can provide a means for cross-pollination and integration across the domains of cognitive neuroscience, working memory, computational modelling, and artificial intelligence (or at least heuristic optimization, for the latter). Both of these parallels give credence to Coherencer as a useful model of certain processes in human cognition. Future research will focus on more advanced thresholds and feature spaces (e.g., spatial relations in addition to co-occurrence), comparisons with other heuristic optimization models, and decreasing the divide between the formalization and research in cognitive neuroscience.

References

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences*, *4*(11), 417-423.

Breault, V., Ouellet, S., Somers, S., & Davies, J. (in press). SOILIE: A computational model of 2D visual imagination. In R. West & T. Stewart (eds.), *Proceedings of the 11th International Conference on Cognitive Modeling*, Ottawa: Carleton University.

Cockbain, J., Vertolli, M. O. & Davies, J. (2013). Creative imagination is stable across technological media: the Spore Creature Creator versus pencil and paper. *The Journal of Creative Behavior*.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87-114.

Davies, J., Atance, C. & Martin Ordas, G. (2011). A framework and open questions on imagination in adults and children. *Imagination, Cognition, and Personality, Special Issue on Mental Imagery in Children*. 31:1-2, 143-157.

Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., & Compte, A. (2009). Mechanism for top-down control of working memory capacity. *Proceedings of the National Academy of Sciences*, *106*(16), 6802-6807.

Eliasmith, C., & Thagard, P. (1997). Waves, particles, and explanatory coherence. *British Journal for the Philosophy of Science*, *48*, 1-19.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in cognitive sciences*, *11*(7), 299-306.

Huang, S. J., Yu, Y., & Zhou, Z. H. (2012). Multi-label hypothesis reuse. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 525-533). ACM.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron*, *63*(6), 889-901.

Hullett, K., & Mateas, M. (2009, April). Scenario generation for emergency rescue training games. In *Proceedings of the 4th International Conference on Foundations of Digital Games* (pp. 99-106). ACM.

Konkel, A., & Cohen, N. J. (2009). Relational memory and the hippocampus: Representations and methods. *Frontiers in Neuroscience*, *3*, 166–174.

Markman, K. D., Klein, W. M., & Suhr, J. A. (Eds.). (2012). *Handbook of imagination and mental simulation*. Psychology Press.

Mullally, S. L., & Maguire, E. A. (2013). Memory, Imagination, and Predicting the Future A Common Brain Mechanism?. *The Neuroscientist*, 19(3), 224-328.

Simon, H. A. (1991). Bounded rationality and organizational learning.*Organization Science*, *2*(1), 125-134.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, *12*, 435-467.

Thagard, P. (1992). Adversarial problem solving: modeling an opponent using explanatory coherence. *Cognitive Science*, *16*, 123-149.

Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MIT Press.

Vertolli, M. O. & Davies, J. (2013). Visual imagination in context: Retrieving a coherent set of labels with Coherencer. In R. West & T. Stewart (eds.), *Proceedings of the 12th International Conference on Cognitive Modeling*, Ottawa: Carleton University.

Vertolli, M. O., Breault, V., Ouellet, S., Somers, S., Gagné, J. & Davies, J. (under review). Theoretical assessment of the SOILIE model of the human imagination, *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Quebec City, QC: Cognitive Science Society.

Von Ahn, L., & Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 319-326). ACM.

Von Ahn, L., Liu, R., & Blum, M. (2006). Peekaboom: a game for locating objects in images. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 55-64). ACM.

Zhang, M., & Zhou, Z. (2013). A Review on Multi-Label Learning Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, (99).