Solving Valid Syllogistic Problems using a Bidirectional Heteroassociative Memory

Marie-France Hébert (mhebe060@uottawa.ca)
School of Psychology, 136 Jean Jacques Lussier, Vanier Hall
Ottawa, ON K1N 6N5 Canada

Sylvain Chartier (sylvain.chartier@uottawa.ca)
School of Psychology, 136 Jean Jacques Lussier, Vanier Hall
Ottawa, ON K1N 6N5 Canada

Christophe Tremblay (ctrem040@uottawa.ca)
School of Psychology, 136 Jean Jacques Lussier, Vanier Hall
Ottawa, ON K1N 6N5 Canada

Abstract

Classical syllogistic reasoning, also known as Aristotelian reasoning, is of particular interest in cognition. Such reasoning, which can seem simple at first, is known to be associated with high error rates. Although some research has been done on this topic, the underlying mechanisms used by human beings remain largely unknown. To understand the underlying cognitive properties associated with solving syllogistic problems, this study uses a connectionist approach composed of three steps inspired from Laird and Bara (1984): spatial representation, associative memory, and alternative searching conclusion. Results show that the network produces similar performances as humans.

Keywords: Syllogistic reasoning; artificial neural network; neurodynamic modeling.

Introduction

Classical syllogistic reasoning has been studied since Aristotle's era. One of his works on deductive reasoning, Prior Analytics, discusses the syllogistic method. Aristotelian reasoning is defined by the act of problem solving solely based on propositions. A syllogism is a type of deductive logic that asks for a verification of the truthfulness of a conclusion based on the presupposed validity of two premises. An example of a syllogism might be: *If all human are mammals and all mammals have four legs, then all humans have four legs*. In this case, the syllogism is valid because the structure of the conclusion given is true. (If all As are Bs and all Bs are Cs, then all As are Cs). Logic reasoning and its cognitive foundations are of particular interest in psychology and neuroscience. Nowadays, the idea of mental models (Johnson-Laird & Bara, 1984) dominates the literature explaining human syllogistic reasoning. Although syllogisms are made of two affirmations, their conclusion is far from being simple. In fact, humans frequently make mistakes as a problem increases in complexity (Dickstein, 1976; Dickstein, 1978; Erickson, 1974). For example, given the problem *No Bs are As and all Bs are Cs*, what is the relationship between As and Cs? In this problem, the possible conclusions are that *some As are Cs*, *all As are Cs*, *no As are Cs*, *some As are not Cs*, *some Cs are As*, *all Cs are As*, *no Cs are As*, or *some Cs are not As*. However, the only valid final conclusion is that *some Cs are not As*, which is not intuitive.

As shown in the previous examples, a syllogism is comprised of two premises. One premise is deemed *major* and the other *minor*, with each premise leading to a conclusion. The major premise is made up of two terms: the conclusion's predicate and the middle term. The minor premise is also made up of two components: the conclusion's subject and the middle term. Notice that the middle term is a component in both premises and links the conclusion's subject and predicate together. In addition, every premise can be formed from one of the following four qualifications: *the universal affirmative*, *the particular affirmative*, *the universal negative*, and *the particular negative*. Every premise comprises of two terms, the antecedent and the consequent. There are 64 syllogistic problems, 27 of which are valid, meaning they have a valid conclusion.

In the field of cognition, there are three principal currents of thought explaining syllogistic reasoning. The first current, formal logic, is based on language rules (Rips, 1994). The second current postulates that the probabilistic heuristics of an event's occurrence are involved in explaining logical reasoning mistakes (Chater & Oaksford, 1999). The last current is based on mental schemes. While performing syllogistic problems, humans use mental representations of the information given (Johnson-Laird & Bara, 1984). In this case, it is more probable that the information is represented by, for example, geometric forms or images instead of mathematical symbols. This study will focus on the last current of thought.

The use of computational systems can be a great tool for a better understanding of syllogistic reasoning. Recent research in syllogisms has focused on modeling properties of syllogisms, in particular on the influence of term order and on the number of representations needed to solve these reasoning tasks. For example, in a study conducted by Bara, Bucciarelli, and Lombardo in 2001, it was hypothesized that the order of terms influences performance. In other words, they support that the information is reorganized in order to simplify the problem by putting the middle term adjacent to each other. The performances of the system on easy to difficult problems were compared to human performances and were shown to correlate. In fact, on three different problems with a range of difficulty levels, the easiest one was the most successful, followed by the intermediate one, and finally by the hardest. Humans seem to organize information in a transitive manner so that the terms within the premises fit within each other. This kind of strategy leads to systematic mistakes in the conclusions provided because the premises are reorganized according to the transitive order of the terms. For example, if the syllogism is in the form B-A-C-B, by inverting the first premise with the second, the result becomes C-B-B-A, which facilitates organizing the information into memory. This study shows that the number of mental representations needed is not the only variable that explains the difficulty in syllogisms; the figure (term order) is another important aspect of the difficulty.

In fact, the difficulty of a syllogism is a function of several factors: the order of the terms in the premises, the number of possible conclusions, the presence of negation, the presence of a quantity proposition, the likelihood that a syllogism will lead to a generalization mistake, and the way the problem can be represented schematically.

The current study aims to determine if a computational system composed of spatial representation of premises combined with a recurrent associative memory neural networks can replicate human performance on syllogistic problems. The simulations will be performed according to the three steps proposed by Johnson-Laird and Bara (1984): 1) the integration and interpretation of the premises, 2) the formulation of a first conclusion based on the representation of the premises, and 3) the search for alternative conclusions. Finally, if the syllogism is judged valid, a general conclusion is generated. An inference will be valid if it is true for every possible interpretation of its premises (Laird & Bara, 1984).

The neural network must be able to generalize to new premise representations of other syllogistic problems. In addition, the information handled by the BHM model needs to be preprocessed in accordance with mental models (Johnson-Laird & Bara, 1984). This preprocessing must be done using a system of spatial representation of the stimuli. The stimuli must not be unique. Thus, several images can be used to represent one problem because syllogistic problem solving is based on the search for alternative conclusions through different possible representations of the premises. In addition, the spatial representation must also take into account the presence of negation, the relative size of the terms, and the term order in the generation of representations. As for the neural network model, it must be able to associate the right conclusions to the premises of a given syllogistic problem.

The objective of the present study is to use topographic maps as a fundamental basis in syllogistic problem solving. These maps will be used as inputs to an associative memory. Results from the simulation will be compared with the results from human participants.

The next section will introduce the idea behind the representation that will be used to encode a syllogistic problem. The bidirectional heteroassociative memory (BHM) neural networks (Chartier & Boukadoum, 2006, 2011) are then presented, followed by the simulation, results, and discussion sections.

Spatial representation of the stimuli

Following the idea that a spatial representation of the information is formed while performing a syllogisms task, the stimuli should be represented as a topographic map. Therefore, they exist multiple ways of spatial representation arrangements. An important element of syllogistic reasoning is the positive or negative character of the quantifier. A premise can be iconically represented. Such iconic representation can be schematized as a diagram (Erikson, 1974). A diagram can visually account for the negative and positive characters of the quantifier. In such a diagram, negative information is visualized outside of another entity, while positive information is visualized as included inside another entity. Based on that type of representation, we hypothesize that negative information is visualized as more peripheral than positive information. In other words, the mental image of negative information would be more fuzzy and far from the person's attentional field, while the mental image of positive information would be situated at (or at least closer to) the center of the person's attentional field. According to this hypothesis, we deduce that it would be easier to remember positive information than negative information, as it would be more central in the person's cognitive representation. Moreover, this may lead to a bigger probability of forgetting about alternative representations of the information. Figure 1 illustrates this idea.

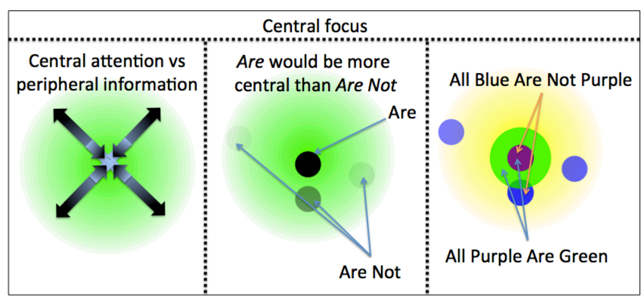

Figure 1: Illustration of the central panel

The left box illustrates the idea that the attention could be focused on the center field or on the peripheral region of a mental scheme. The center box shows that a positive quantifier would be more central than a negative quantifier. The right box illustrates an example where it is easy to forget about an alternative representation if the negative information remains far from the center of the cognitive attentional field. In this example, it would be less probable to think about a blue circle that touches the green circle.

Model: The bidirectional Heteroassociative Memory (BHM)

A premise in a syllogism can be represented in several ways and the model must be robust to this constraint. This constraint should be taken into account by a Bidirectional Heteroassociative Memory (BHM) (Chartier & Boukadoum, 2006, 2011). Using a BHM provides the system with a more realistic process, since memory is an important part of syllogistic reasoning (Gilhooly, Logie, & Wynn, 1999). For example, interference in short-term memory may lead to increase the resolution difficulty due to an overload of information in the brain. The lack of fluid memory may lead to particular errors, particularly when the premises are formulated in a way where humans need to inverse spatial information throughout the whole reasoning process. Moreover, the BHM has the particularity to deal with noise, allowing the use of multiple ways to present the same premise to the system. In order to solve syllogisms, generally more than one spatial representation of the same linguistic information is needed. The BHM is a neural network that is most likely able to deal with this constraint. Following these arguments, the BHM is a great way to provide a more realistic overall system to study syllogistic reasoning.

Architecture of the BHM

The architecture is made of two Hopfield-like (Hopfield, 1982) neural networks interconnected in a head-to-toe fashion, which allows the association of a pair of stimuli. The architecture is presented in Figure 2, where **x**(0) and **y**(0) represent the initial state of the input vectors to be associated, t represents the number of iterations cycle performed, and **W** and **V** are the weight connections.

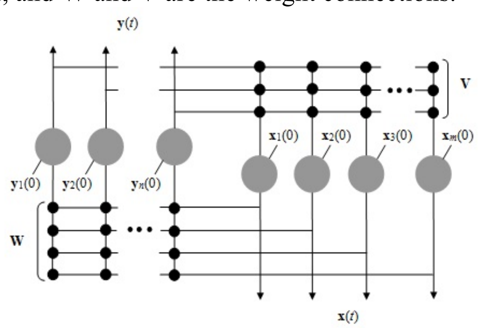

Figure 2: Architecture of the BAM

Transmission function

The transmission is defined by the following equations:

$$\forall i, \ldots, N, \mathrm{y}_{i(t+1)} = f\left(\mathrm{a}_{i(t)}\right) = \begin{cases} 1, & \text{if } \mathrm{a}_{i(t)} > 1 \\ -1, & \text{if } \mathrm{a}_{i(t)} < -1 \\ (\delta+1)a_{i(t)} - \delta a_{i(t)}^3, & \text{else} \end{cases} \tag{1a}$$

and

$$\forall i, \ldots, M, \mathrm{x}_{i(t+1)} = f\left(\mathrm{b}_{i(t)}\right) = \begin{cases} 1, & \text{if } \mathrm{b}_{i(t)} > 1 \\ -1, & \text{if } \mathrm{b}_{i(t)} < -1 \\ (\delta+1)\mathrm{b}_{i(t)} - \delta \mathrm{b}_{i(t)}^3, & \text{else} \end{cases} \tag{1b}$$

where N and M are the number of units in each layer, i is the unit index, **y**(t+1) and **x**(t+1) are the output given at time t+1, δ is a general transmission parameter, and **a** and **b** are the activation. These activations are obtained the usual way: **a**(t)=**Wx**(t) and **b**(t)=**Vy**(t).

Learning rule

The weight connections are modified following Hebbian/Anti-Hebbian principles (Storkey & Valabregue, 1999; Bégin & Proulx, 1996):

$$\begin{aligned} \mathbf{W}_{(k+1)} &= \mathbf{W}_{(k)} + \eta\left(\mathbf{y}_{(0)} - \mathbf{y}_{(t)}\right)\left(\mathbf{x}_{(0)} + \mathbf{x}_{(t)}\right)^{\mathrm{T}} \\ \mathbf{V}_{(k+1)} &= \mathbf{V}_{(k)} + \eta\left(\mathbf{x}_{(0)} - \mathbf{x}_{(t)}\right)\left(\mathbf{y}_{(0)} + \mathbf{y}_{(t)}\right)^{\mathrm{T}} \end{aligned} \tag{2}$$

where k is the trial number, **V** and **W** are the weight connections, and h is a small positive learning parameter. The weight connections are adjusted in function of the difference between the initial activation state (**y**(0) and **x**(0)) and the output at time t (**y**(t) and **x**(t)). The network converges to a solution when **x**(0) = **x**(t) or when **y**(0) = **y**(t). In other words, the weights converge when the difference between the desired value and the actual time value is null.

Simulation

Three-step process

The proposed system consists of a succession of the three-step process proposed by the upholders of mental models. The first step is the integration and the interpretation of the premises. This was based on the postulate that human beings integrate and interpret syllogistic premises as mental schemes. At this step in the system, the information on each problem is preprocessed in function of predefined rules before being transformed into input vectors for the BHM. The premises are expressed graphically and are juxtaposed to form one input pattern. The BHM then gives the associated conclusion (output). Finally, the search for alternative conclusions is accomplished for a fixed number of trials. A different representation is thus given as an input in order to see if the BHM will generate the same or a different conclusion. The more alternative possible representations are allowed, the higher the probability of finding the correct answer. If the conclusions given by the BHM are contradictory, the syllogistic problem will not be considered valid. If a conclusion is still true regardless of its representation, the problem will be deemed valid.

Simulations were performed in order to compare the system with human performance. Every simulation followed the three steps described previously and was performed on one possible representation of the 27 valid problems. Therefore, the BHM had to learn only 27 pairs of patterns rather than every possible premise (3884) to a respective conclusion.

Methodology

Following the topographic rules illustrated in Figure 3, colored squares are used to illustrate the three terms (major term, middle term, and minor term). The figure also illustrates the rules used to form a stimulus that will be shown to the network. In order to represent quantities, whether *all* or *some*, two sizes are used: four pixels for *all* and two pixels for *some* (upper right corner of Figure 3). The information is positioned in the center of the area restricted to the antecedent or outside of this area to represent the affirmative or negative of the antecedent, respectively. Thus the spatial information represents the quality (*is* or *is not*) and the size of the square represents the quantity.

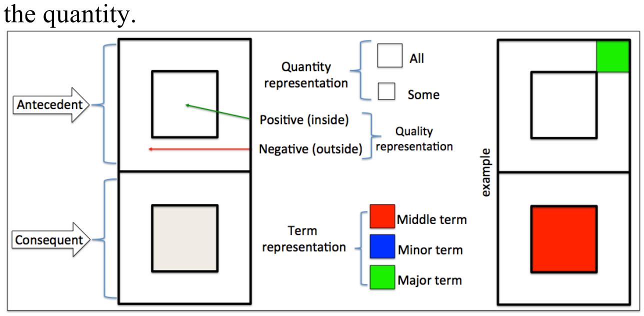

Figure 3: Representation system for the premises and at the right the following premise example: All Greens are not Red

The square representing the quantity of the antecedent is the same color as the term. If the antecedent is the middle term, the color is red; for the minor term, the color is blue; and for the major term, the color is green. Finally, if the consequent is the middle term, the color of the area restricted to the consequent is red, it is blue if the consequent is the minor term, and green if it is the major term. The following example can be seen at the right side of Figure 3: All greens are not red. In this example, since the quantity of the antecedent is *all*, the square is bigger. In addition, since the quality is negative, the square is situated on the exterior to the area restricted to the antecedent.

A list of images representing every possibility for the first premise (four images) for every valid syllogistic problem is established. This list is built from the different possibilities issued from the construction rules as previously illustrated (Figure 3). The first list of stimuli contains the major premises, made up of the color green for the major term and red for the middle term. Another list of images representing every possibility for the second premise (four images) for every valid syllogism is established using the same process. This second list then holds the minor premise, made up of the color blue for the minor term and red for the middle term. For every problem, there are 16 possible representations. Since there are 27 valid syllogisms, there are a total of 432 syllogistic problem representations. To limit the number of images, the information representing the antecedent is positioned in the corners of the area restricted to the antecedent (upper left, upper right, lower left, lower right). For this reason, there are only four possible images per premise.

Every premise is made of 9 X 18 pixels, for which three values give a color pixel. These dimensions were chosen as they permit the smallest representation that allows for an accounting of all the characteristics needed for a representation. The vector (-1,1,1) defines the red color, (1,-1,1) defines green, and (1,1,-1) blue. The correlation between each pair of colors was -0.5. The correlations between the stimuli vary between 0.21 and 0.996. The correlations can be high because the quantifier *some* is represented by only two pixels, so sometimes two images can differ by only few pixels. A first juxtaposition of the two premises of a problem forms the problem's image, which is then vectorized. This vector must be associated with another vector that represents the conclusion. This associated conclusion vector represents a connection between the subject (minor term) and the predicate (major term). These input vectors have a dimensionality of 972 pixels (2 X 9 X 18 X 3).

The conclusion is built according to the following: Firstly, for every premise, the consequents must be situated minimally on the same topographic region as the antecedents. Secondly, the size of the consequent can vary and be larger than the antecedent. The consequent on the conclusion is limited to three different sizes, thereby making nine possible conclusions for one given representation (remember that the problem is made up of two premises). The training is thus accomplished under variability for each syllogism representation. For a given syllogistic problem, the BHM must associate an answer representing the connection between the subject (minor term) and the conclusion's predicate (major term). Considering every possible valid representation for a syllogistic vector, using the cartographic rules, and considering the possible spatial representations for the conclusions, 3888 pairs of stimuli are possible. Of those possibilities, the BHM model associates a subsample of 27 pairs: one pair per valid syllogism.

Learning was deemed accomplished when the sum of squared error was less than 10^{-4}. Usually, learning required approximately 500 epochs.

Following the learning, recall tests were performed. During a recall trail, random selections of the stimuli representing the two premises from the whole sample (3888 pairs) were

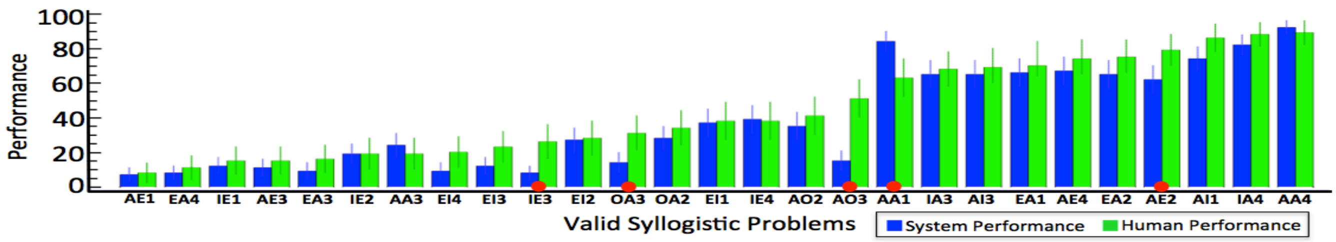

Figure 5: Observed (blue bar) and predicted (green bar) proportions of success on the 27 valid syllogisms. Premises of the syllogisms problem are abbreviated using A for *all*, E for *none*, I for *some*, and O for *some are not*. Numbers 1 to 4 represent the syllogistic figure as described in the introduction. Error bars show 95% confidence intervals. Red circles indicate when the difference does not fall into the confidence interval.

randomly selected. An output (conclusion) is generated using the BHM according to Equation 1. If an alternative conclusion is allowed, a novel stimuli representation is selected and a conclusion is generated. The number of alternative conclusions followed the given rules: 0 alternative conclusions 90% of the time, 1 alternative conclusion 5% of the time, and 2 alternatives conclusions for 5% as well. The final conclusion is that which is true for all the alternative conclusions. If the conclusions are contradictory for one problem, then the problem is deemed invalid. It is important to note that in some cases the conclusions aren't always similar but are not contradictory. In those cases, the answer is the conclusion that is most conservative. For example, if two alternative conclusions are *some greens are blues* and a third conclusion is *all greens are blues*, the final answer will be *some greens are blues* because for the three alternative conclusions, there are at least some greens that are blues. If the answer is possible in both orders, one of the orders is picked randomly.

The 27 syllogistic problems are tested one after the other. The system is successful if the conclusion given by the system is true. Otherwise, the trial is considered a failure. Every problem is tested 150 times in order to obtain an average value of the performances.

Results

The average performance of the system for every syllogistic problem is calculated in function of the three steps, as described previously for the 27 valid syllogisms. Results illustrated in Figure 5 show both human and system performance. The human performances are taken from the study of Johnson-Laird and Byrne (1991). The correlation between the two is strong, $r(52) = .92$, $p < .01$. Differences that lay outside the 95% confidence interval are marked by a red circle. Another way to look at the performance results is by using an ordinal pattern analysis. First, the 27 syllogisms are ranked from the easiest to the most difficult (Thorngate, 2013) according to human performance. The performance of the system is also ranked in a similar fashion. A total of 351 possibilities of matches can be computed by calculating the number of possible pairs that can be obtained with 27 problems. The 27th problem performance for the system is then supposed to be higher then all the other problems and so on (27 > 26, 27 > 25, ... 27 > 1, 26 > 25, etc.). In order to test how well the predictions match the observed ranking, the proportion of good matches is calculated. A total of 301 good matches on 351 were found, leading to a proportion of .86 ($p < .001$). This proportion is far from .50, the proportion that would have been obtained by chance.

Discussion and Conclusion

The results show that the BHM network does not need to be trained on all possible representation in order to be efficient. For a given simulation, the learning phase is accomplished on a random set of 27 pairs of stimuli, with one representation per syllogistic problem. During recall, novel pairs were presented, which affected the performance of the network. Because BHM develops attractors, its learning can be generalized. This difficulty induced in the network creates variability in the performance. Some syllogistic problems are less affected by stimuli variability. Of course, generalization will be influenced by the correlation. The higher the correlation of the representation, the better the generalization. For example, in *all greens are red* and *all reds are blue*, the possible representations resemble very closely, which will lead to similar conclusions. Similarly, the lower the correlation between the representations, the less likely the BHM is to reproduce the right conclusions. In short, the difficulty level associated with syllogisms could arise from the dissimilarities in their corresponding representations. Another source of variability in performance is the number of allowed alternative conclusions. The network was not able to reproduce human performance on all syllogism problems. For example, the performances on the problems IE3, OA3, and AO3 were significantly different from those of humans. This can be explained by the lack of possible sizes for the consequents of the premises that are built with the topographic rules. Another difference can be observed on problem AA1, where the system performance is much higher than the human performance. This can be explained by the fact that the system does not differentiate the order in which the two premises are presented. In fact, the problem AA1 becomes very easy to the human by simply reversing the two premises. This inversion can alleviate memory loading.

Future empirical study will look at how human performances are influenced by the representation method introduced in the current study without limited time to accomplish the task. Also, future simulations should be done using more than three sizes for the consequents, as was used in this study. Even if some premise representations are less probable than others, there are some alternative conclusions that are still not formulated by the system. Moreover, a more thorough exploration of the parameter of the system should be studied in order to determine its robustness. For example, it might be more probably that the size of the consequent for premises built by humans is closer to the size of the antecedent quantity. Following the idea of Khemlani, Trafton, and Johnson-Laird (2013), using a Poisson distribution for the size of the consequent would be interesting. Finally, it would be interesting to test the system on all syllogisms, valid or not. Multiple series of syllogism could also easily be tested with this system. In fact, when premises are added a more complex syllogism is created. It would then be interesting to simulate a complex case of syllogism for comparison with human performance.

In conclusion, the three-step system of integrating a BHM network allows a basic framework for the study of syllogisms. This provides the system with a realistic human cognitive perspective. The system can then be used to evaluate the performance under various constraints.

Acknowledgements

This research was funded in part by the Fonds Québécois de la Recherche sur la Nature et les Technologies and the Natural Sciences and Engeneering Research Council of Canada. The authors would like to thank Kaia Myers-Stewart, Indesh Singh and Anna Lee-Popham for their great comments.

References

Bara, B. G., Bucciarelli, M., & Lombardo, V. (2001). Model theory of deduction: A unified computational approach. *Cognitive Science*, *25*, 839-901.

Bégin, J., & Proulx, R. (1996). Categorization in unsupervised neural networks: The eidos model. *IEEE Transactions on Neural Networks*, *7*(1), 147-154.

Chartier, S., & Boukadoum, M. (2006). A bidirectional heteroassociative memory for binary and grey-level patterns. *IEEE Transactions on Neural Networks*, *17*(2), 385-396.

Chartier, S., & Boukadoum, M. (2011). Encoding static and temporal patterns with a bidirectional heteroassociative memory. *Journal of Applied Mathematics*, *2011*, 1-34.

Chater, N. & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, *38*, 191-258.

Dickstein, L. S. (1976). Differential difficulty of catagorical syllogisms. *Bulletin of the Psychonomic Society*, *8*(4), 330-332.

Dickstein, L. S. (1978). Error processes in syllogistic reasoning. *Memory & Cognition*, *6*(5), 537-543.

Erickson, J. R. (1974). A set analysis theory of behavior in formal syllogistic reasoning tasks. In R.L. Solso (Ed.). *Theories in cognitive psychology: The Loyola symposium on cognition, 2, 305-330.*

Gilhooly, K. J., Logie, R. H., & Wynn, V. (1999). Syllogistic reasoning tasks, working memory, and skill. *European Journal of Cognitive Psychology*, *11*(4), 473-498.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554-2558.

Johnson-Laird, P. N. & Bara, B. G. (1984). Syllogistic inference. *Cognition*, 1*6*, 1-61.

Johnson-Laird, P. N., & Byrne, R. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Khemlani, S., Trafton, J.G., & Johnson-Laird, P. N. (2013). Deduction as stochastic simulation. *The 12th International Conference on Cognitive Modelling.* Ottawa, Canada: Carlton University.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, MA: MIT Press.

Storkey, J., & Valabregue, R. (1999). The basins of attraction of a new Hopfield learning rule. *Neural Networks, 12*(6), 869-876.

Thorngate, W. (2013). Measuring simulation-observation fit: An introduction to ordinal pattern analysis. *Journal of Artificial Societies and Social Simulation,* 16(2), 4.