Data Acquisition Dynamics and Hypothesis Generation

Nicholas D. Lange (ndlange@gmail.com) †‡
Rick P. Thomas (Rickey.P.Thomas-1@ou.edu) ‡
Eddy J. Davelaar (e.davelaar@bbk.ac.uk) †
†Birkbeck, University of London, Department of Psychological Sciences,
Malet Street, London, UK, WC1E 7HX
‡University of Oklahoma, Department of Psychology
Lindsey Street, Norman, OK, USA 73019

Abstract

When formulating explanations for the events we witness in the world temporal dynamics govern the hypotheses we generate. In our view, temporal dynamics influence beliefs over three stages: data acquisition, hypothesis generation, and hypothesis maintenance and updating. This paper presents experimental and computational evidence for the influence of temporal dynamics on hypothesis generation through dynamic working memory processes during data acquisition. Results suggest that data acquired from the environment undergo dynamic competition in working memory, the results of which dictate the weights allocated to individual data in the generation process.

Keywords: hypothesis generation, temporal dynamics, working memory, abduction, diagnostic reasoning

Introduction

Hypothesis generation is a pre-decisional process by which we formulate explanations and beliefs regarding the occurrences we observe in our environment. The hypotheses we generate from long-term memory bring structure to many of the ill-structured decision making tasks we encounter on a daily basis. As such, hypothesis generation represents one of our most fundamental and ubiquitous cognitive faculties. Given such regularity, it is no surprise that hypothesis generation forms a core component of several professions. Auditors, for instance, must generate hypotheses regarding abnormal financial patterns and mechanics must generate hypotheses concerning car problems. Perhaps the clearest example, however, is that of medical diagnosis. A physician observes a pattern of symptoms presented by a patient (i.e., data) and uses this information to generate likely diagnoses (i.e., hypotheses) in an effort to explain the patient's current disease state. Given these examples, the importance of developing a full understanding of the processes underlying hypothesis generation is clear, as the consequences of impoverished or inaccurate hypothesis generation can be injurious.

When engaged in hypothesis generation tasks, cognitive limitations place constraints on the acquisition of bits of data used to cue long-term memory for the retrieval of likely hypotheses. Important to the present work is the fact that data acquisition most often occurs serially. This, in turn, dictates that individual pieces of data are acquired in some temporal relation to one another. These constraints, individual data acquisition over time and the relative ordering of data, are likely to have significant consequences for hypothesis generation processes. Given these basic constraints it is intuitive that temporal dynamics must form an integral part of any comprehensive account of hypothesis generation processes. In our view temporal dynamics influence beliefs over three stages: data acquisition, hypothesis generation, and hypothesis maintenance and updating (as further data is acquired or judgments and decisions rendered). This paper concerns the temporal dynamics unfolding over the initial data acquisition phase which until now has remained unaddressed.

At present there exists limited data concerning the temporal dynamics of hypothesis generation tasks. Thus, the influences of the constraints operating over these processes are not yet well understood. Until such influences are addressed at an empirical and theoretical level a full understanding of hypothesis generation processes will remain speculative. Interest in understanding these underlying temporal dynamics is increasing however. For instance, Sprenger & Dougherty (2011) found a general recency bias in hypothesis generation whereby people tended to generate hypotheses more consistent with data received later than data received earlier. Additionally, Mehlhorn et al. (2011) investigated how hypotheses' memory activations are influenced by the amount of data that has been received at various time steps finding increases in memory activation with increases in supporting data.

HyGene (Dougherty, Thomas, & Lange, 2010; Thomas, Dougherty, Harbison, & Sprenger, 2008), short for **hy**pothesis **gene**ration, is a computational architecture addressing hypothesis generation, evaluation, and testing. This framework has provided a useful account through which to understand the cognitive mechanisms underlying these processes. Here we extend this work by incorporating working memory dynamics from the context activation model of list memory (Davelaar, et al., 2005) to account for data acquisition dynamics subserving the cued recall process inherent in hypothesis generation.

HyGene & Temporal Dynamics

HyGene rests upon three core principles. First, it is assumed that hypothesis generation represents a generalized case of cued recall. Data observed in the environment (D_{obs}), which one would like to explain, act as cues prompting the retrieval of hypotheses from long-term memory (LTM). For

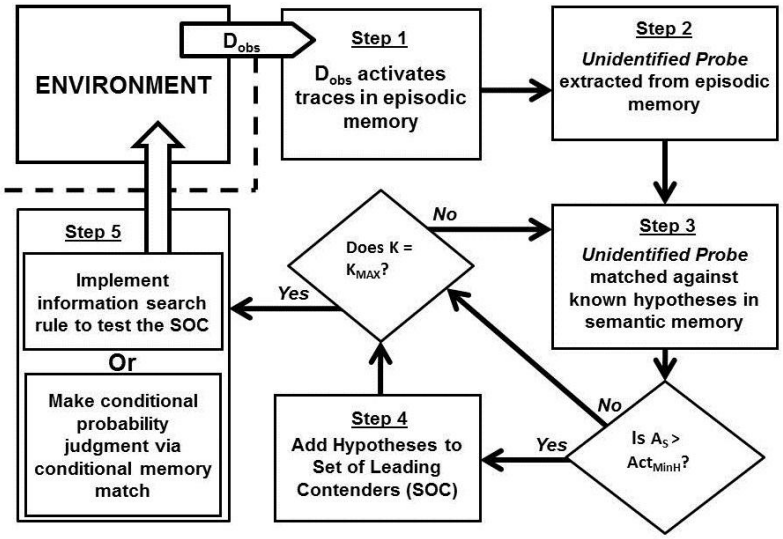

Figure 1: Flow diagram of processing in HyGene

instance, when a physician examines a patient, he/she uses the symptoms expressed by the patient as cues to related experiences stored in LTM. These cues activate a subset of related memories in episodic memory which guide the generation of hypotheses from semantic memory. These retrieval processes are indicated in steps one, two, and three of Figure 1. As viable hypotheses are retrieved from LTM they are placed in the Set of Leading Contenders (SOC) as demonstrated in step four. The SOC represents HyGene's working memory construct to which the second principle applies.

The second principle holds that the quantity of hypotheses that can be maintained at one time is constrained by cognitive limitations as well as task characteristics. That is, the more working memory resources that one has available to devote to the generation and maintenance of hypotheses, the more accommodating the SOC will be of additional hypotheses. Working memory capacity places an upper bound on the amount of hypotheses (and data) that one will be able to maintain at any point in time. In many circumstances, however, attention will be divided by a secondary task. Under such conditions this upper bound is reduced as the alternative task siphons resource that would otherwise allow the population of the SOC to its unencumbered capacity (Dougherty & Hunter, 2003a; Dougherty & Hunter, 2003b; Sprenger & Dougherty, 2006; Sprenger et. al., 2011).

The third principle states that the hypotheses maintained in the SOC form the basis from which probability judgments are derived and provide the frame from which hypothesis testing is implemented. This principle underscores the function of hypothesis generation as a pre-decisional process underlying higher-level decision making tasks and can be seen as step five in the diagram.

These assumptions form the core of HyGene's theoretical framework. HyGene in its current form is *static* with regards to data acquisition and utilization. The model receives all available data from the environment simultaneously and engages in only a single iteration of hypothesis generation. Given the static nature of the model, each piece of data used to cue LTM contributes equally to the recall process. There is reason to suspect, however, that all available data do not generally contribute equally. What is needed is an understanding of working memory dynamics as data acquisition, hypothesis generation, and maintenance processes unfold and evolve over time in hypothesis generation tasks.

A Dynamic Model of Data Acquisition and Hypothesis Generation

We now forward a dynamic version of HyGene in which the activations of individual pieces of data acquired from the environment fluctuate over time in working memory prior to hypotheses being generated from long-term memory. The activation levels possessed by each piece of data at the time of generation are used as weights in the retrieval of hypotheses. This allows the activation of each piece of data in working memory to govern its individual contribution to the generation process. These dynamic working memory processes were borrowed from the context-activation model of memory (Davelaar et al., 2005). This model dictates that the activations of the items in working memory systematically fluctuate over time as the result of competing processes described by Equation 1.

$$x_i(t+1) = \lambda x_i(t) + (1-\lambda)\{\alpha F[x_i(t)] + I_i(t) - \beta \sum_{j \neq i} F[x_j(t)] + N(0,\sigma)\}$$

Equation 1: activation calculation of the context-activation model

The activation level of each item in the buffer, x_i, is determined by the items activation on the previous time step, self-recurrent excitation that each item recycles onto itself α, sensory input I, inhibition from the other active items β, and zero-mean Gaussian noise ξ with standard deviation σ. λ is the Euler integration constant that discretizes the differential equation.

Figure 2 illustrates the interplay between these competing forces in noiseless runs of the buffer when five pieces of data have been presented to the model for a fast rate of 100 iterations (top panel) and for a slower rate of 1500 iterations (bottom panel). The activation of each data rises as it is presented to the model and its bottom-up sensory input contributes to the activation. These activations are then dampened in the absence of bottom-up input as inhibition from the other items drive activation down. Self-recurrency can keep an item in the buffer in the absence of bottom-up input, but this ability is in proportion to the amount of competition from other items in the buffer. As can be seen, the fast presentation rate, in comparison to the slow rate, results in less competition from later items as the truncation of sensory input renders them less competitive. Importantly, this shift from recency to primacy with increasing presentation rate is a unique prediction made by this dynamic buffer and challenges other buffer models (Davelaar, et al., 2005).

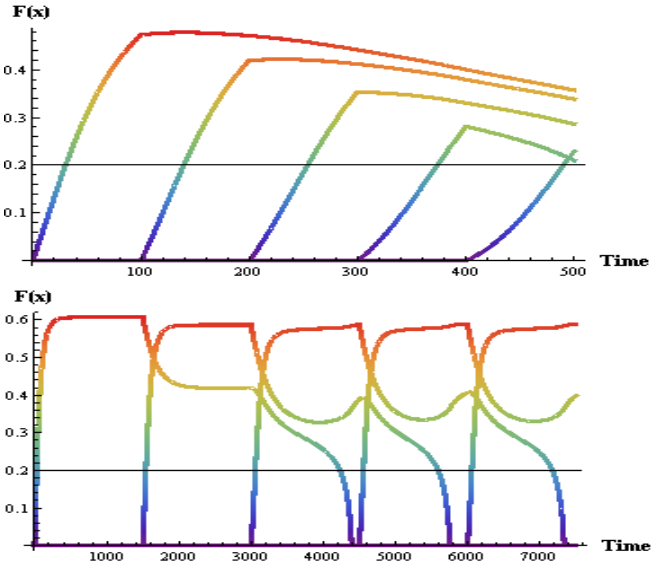

Figure 2: Activation trajectories for 5 sequentially received data at fast presentation rate (top) and slow presentation rate (bottom)

HyGene utilizes a representation from the multiple trace global matching models of MINERVA II (Hintzman, 1986, 1988) and the decision making model MINERVA-DM (Dougherty et al., 1999)[1]. Separate episodic and semantic memory stores are present in the model. While semantic memory stores only individual prototypes of each disease, each experience the model acquires is represented in episodic LTM as a series of concatenated minivectors of 1s, 0s, & -1s where each minivector represents a hypothesis or data. That is, each trace is made up of one hypothesis and several pieces of data (in our case four). Retrieval is initiated when D_{obs} are matched against the data minivectors in LTM. This results in an activation level for each trace where a greater overlap in features present in the trace and in the D_{obs} results in greater activation. The weightings from the data acquisition buffer are used to weight the activations of each minivector in episodic memory at this point in retrieval. Therefore, the activation levels associated with each trace are directly influenced by the weightings for each data supplied by the dynamic working memory processes of the buffer.

Once these activation values have been obtained, only a subset of the episodic traces activated over a criterion are used for further processing in the model. From this subset of traces a probe is derived as a cue to semantic memory for the generation of hypotheses. This cue is matched against all known hypotheses in semantic memory. The activation values for each hypothesis serve as input into sampling via Luce's choice rule. Generation proceeds until a stopping rule is reached based on the total number of resamplings of previously generated hypotheses (i.e., retrieval failures).

We now present two experiments investigating separate consequences of hypothesis generation being extended over time. The first experiment examines how the mere serial position of a diagnostic datum influences the generation of the hypothesis it implies. Experiment two examines how processing time (i.e., presentation duration) per datum influences the contributions of the individual data in the generation process. The novel model of dynamic data acquisition and hypothesis generation discussed above is used to simulate the findings from both experiments. Critically, although many instantiations of a working memory buffer may predict the results from Experiment 1, the results from Experiment 2 provide support for our specific buffer instantiation, as borrowed from the context-activation model, underlying data acquisition in hypothesis generation tasks.

The Influence of Data Acquisition Dynamics on Hypothesis Generation

Experiment 1

The generalized order effect paradigm was developed by Anderson (1973) to examine the differential weighting of descriptive attributes presented in impression formation tasks. The procedure involved embedding a fixed list of information with a critical piece of information at various serial positions thereby allowing differences in the final rating to be uniquely attributable to the serial position of the critical data. The present experiment represents an

[1] For a more thorough treatment of HyGene's computational architecture please see Thomas, Dougherty, Harbison, & Sprenger, (2008) or Dougherty, Thomas, & Lange (2010)

adaptation of this paradigm to a simulated medical diagnosis task to assess the impact of specific data serial positions on hypothesis generation.

Method

Participants Seventy-two participants participated in this experiment for course credit.

Design The design of Experiment 1 was a one-way between-subjects design with data order as the independent variable. The ecology for this experiment as defined by the conditional probabilities between the hypotheses and data is shown in Table 1. Each of the values appearing in this table represents the probability that the symptom will be present (e.g., fever) given a particular hypothesis whereas the complementary probability represents the probability of the symptom absence. As demonstrated in the table, the only diagnostic piece of data was D1 whereas the remaining cues, D2-D4, were non-diagnostic.

Table 1: Disease (Hypothesis) x Symptom (Data) ecology of Experiment 1.

		Symptoms			
		D1	**D2**	**D3**	**D4**
Diseases	**H1: Metalytis**	0.8	0.6	0.6	0.6
	H2: Zymosis	0.2	0.6	0.6	0.6
	H3: Gwaronia	0.2	0.6	0.6	0.6

Table 2 displays the four data orders. Each of these orders was identical (D2 → D3 → D4) except for the position of the D1 data within them.

Table 2: Data presentation orders.

	→ Presentation Position →			
	1	**2**	**3**	**4**
Order 1	**D1**	D2	D3	D4
Order 2	D2	**D1**	D3	D4
Order 3	D2	D3	**D1**	D4
Order 4	D2	D3	D4	**D1**

Procedure The procedure was comprised of two stages. The first stage was an exemplar training task in which a series of hypothetical pre-diagnosed patients was presented to the participant in order for them to learn the contingencies between the hypotheses and data through repeated experience. Each of these patients was represented by a diagnosis at the top of the screen (H1, H2, or H3) and a series of test results (i.e., symptoms) pertaining to the columns of D1, D2, D3, and D4. Over the course of the training phase the specific test results precisely respected the disease-symptom contingencies appearing in Table 1.

Following an arithmetic distraction task, the second stage of the procedure commenced. This was an elicitation phase in which we implemented our manipulation of data order and assessed hypothesis generation performance. The participants were then told that they were now going to see an individual patient's symptoms and would then be asked to report the most likely diagnosis for the patient. The participant triggered the onset of the patient's data stream at their readiness. Each datum of was presented individually for 1.5 seconds. The order in which the data were presented was determined by the order conditions as shown in Table 2. Following the presentation of the last datum the participant responded with the most likely disease.

Results

Empirical Nominal logistic regression was carried out on the generation data to examine the effect of data serial position on the generation of H1 (Metalytis), the disease with the greatest posterior probability given the data. There was a significant trend for H1 being reported as the most likely hypothesis as the serial position of the diagnostic data increased, $\chi^2(1) = 4.32$, $p < 0.05$.

Computational To simulate Experiment 1, the model's episodic memory was endowed with the Hypothesis-Data contingencies described in Table 1. On each trial each piece of data was presented to the buffer for 1500 iterations (mapping onto the presentation duration of 1500 ms) and the order of the data was manipulated to match the data orders used in the experiment. 1000 iterations of the entire simulation were run for each condition[2]. The model data was supplemented with a constant guessing parameter of 0.31 across all conditions.

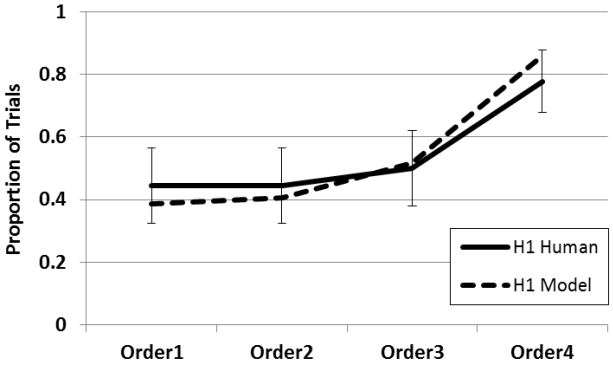

Figure 3: Human and Model results for Experiment 1 plotting the probability of reporting H1 as most likely across order conditions. Error bars represent standard errors.

As is demonstrated in Figure 3, the model is able to capture the empirical data quite well. This effect is directly attributable to the weights from the buffer being applied to the generation process.

[2] The parameters used for this simulation were the following. HyGene: L=0.85, Ac=0.075, Phi=4, KMAX=8 Buffer: Alpha=2.0, Beta=0.2, Lambda=0.98, Delta=1.0

Experiment 2

Method

Participants One hundred and twenty four participants participated in this experiment for course credit.

Design The design of Experiment 2 was a one-way between-subjects design with the presentation rate of the data as the independent variable. The ecology for this experiment appears in Table 3. The important aspect of this ecology is that the early data (D1 and D2) are diagnostic in favor of H1 whereas the later data (D4 and D5) favors H2.

Table 3: Disease (Hypothesis) x Symptom (Data) ecology of Experiment 2.

	D1	**D2**	**D3**	**D4**	**D5**
H1:Metalytis	0.8	0.7	0.5	0.3	0.3
H2:Zymosis	0.3	0.3	0.5	0.7	0.8

Procedure The procedure of this experiment was very similar to that of Experiment 1. Participants learned the hypothesis-data contingencies in an exemplar training phase prior to elicitation. However, in between these two phases of the experiment there was a learning test to discriminate participants that had learned the contingencies well from those that did not learn the contingencies. For this test the participants were provided with an individual piece of data and asked what the most likely hypothesis was. Their total learning score was the amount of correct responses in this task[3].

In the elicitation phase the participants were provided with the data in the order in which they appear in Table 3, that is, consecutively from D1 to D5. Directly following the last piece of data the participant entered the disease they thought was most likely given the patient's symptoms. What varied between participants was the rate at which these data were presented. Half of the participants were presented the data at a fast rate (144 ms each) while the other half were presented the data at a slow rate (1504 ms each).

As displayed in Figure 2, the context-activation model predicts the fast presentation rate to lead to the earlier data residing more strongly in working memory following D5 whereas the model predicts the opposite for the slower presentation rate. Therefore we predicted that the fast presentation rate should lead to greater relative activations of early data thereby leading to greater generation of H1, whereas the opposite would be the case when the data are presented slowly leading to a preference for H2 and accordingly a lower rate of H1 relative to the fast condition.

Results

Empirical Although the rate of H1 selection was slightly higher in the fast presentation rate condition, this difference did not reach significance, $z = 1.27$, $p = 0.102$. A further analysis was performed within groups of high learning and low learning participants based on their performance in the learning test. Those scoring higher than 60% were counted as high learners and those scoring lower were counted as low learners. Conditional analyses within each learning group revealed a marginal effect of presentation rate for the low learners, $z = 1.6$, $p = 0.054$ and no effect for the high learners, $z = 0.34$, $p = 0.367$. This result reflects the fact that the trend witnessed in the overall data was, somewhat counter-intuitively, due to those that did not learn the contingencies in the task as fully. We explain this effect below with our model.

Computational To simulate Experiment 2 the model was endowed with experience in the ecology displayed in Table 3. The manipulation of presentation rate was implemented in the model by varying the number of iterations the model was presented each piece of data. For instance, in the fast condition each piece of data received bottom-up input for 100 iterations whereas in the slow condition each piece of data received bottom up activation for 1500 iterations.

In line with the empirical result, however, we are not solely concerned with capturing differences in presentation rate, but we are additionally interested in capturing the difference that manifesting between the high and low learning groups. We posit this difference to be attributable to the role of working memory capacity (WMC). It is likely that high capacity participants were better able to learn the contingencies as each exemplar provided several bits of information for encoding. Furthermore, successful learning likely included some form of hypothesis testing carried out over successive exemplars which would be cognitively taxing, but beneficial to learning. Therefore we suggest that the high learning group possessed a greater proportion of high capacity participants.

In the present analysis, we ask if differences in a parameter governing the emergent capacity of the buffer could explain the presence of a presentation rate effect amongst low learners and its amelioration amongst high learners. As the beta parameter governs the strength of the global inhibition that is applied to each item this parameter can be used to impose capacity constraints (Davelaar, 2007). As beta is increased, competition between items is increased and fewer items will cohabitate the buffer. We manipulated beta at two levels to capture low learning/low capacity (beta= 0.1) and high learning/capacity (beta=0.05) and used presentation rates of 100 iterations (fast rate) and 1500 iterations (slow rate)[4]. This resulted in averaged summed activations in the buffer of 2.22 in the slow rate and 2.33 in the fast rate under beta=0.05 and values of 1.96 in the slow rate and 1.74 in the fast rate under beta=0.1. Therefore, more activation was present in working memory on average when beta=0.05 relative to beta=0.1. This entire simulation was run for 700 model runs of each condition.

[3] Responses were counted automatically correct for responses to the D3 data as both hypotheses were equally likely.

[4] All other model parameters were the same as those used for Experiment 1.

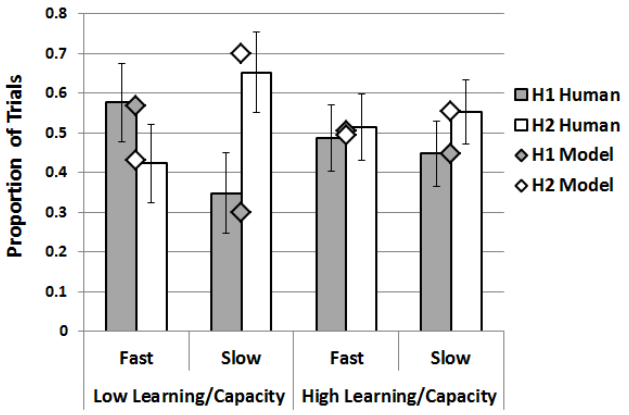

Figure 5: Human and Model results for Experiment 2 plotting the probability of generating H1 and H2 as most likely by presentation rate and learning/capacity groups. Error bars represent standard errors.

As demonstrated in Figure 5, the model is able to capture the patterns in the data occurring with differences in presentation rate and differences in learning between participants.

Discussion

We presented and tested an extension of the HyGene model. The model was endowed with a dynamic working memory buffer and adequately captured a recency bias in generation (Experiment 1). In addition, the sensitivity of the dynamic buffer to presentation rate was shown to influence hypothesis generation (Experiment 2). Moreover, individual differences in learning or WMC interacted with the balance of incorporating primacy and recency items in the decision. Moreover, the ability of our model to capture the results from both experiments lends credence to our specific buffer implementation.

The present work demonstrates the utility of understanding working memory dynamics during *data acquisition* (cf. Mehlhorn et al., 2011) and suggests that the activations of individual pieces of data in working memory govern their individual contributions to the hypothesis generation process. The model presented here will be extended in future work such that the activations of hypotheses themselves will be subject to the competitive buffer dynamics demonstrated here. This model will address the hypothesis maintenance and updating components of temporally dynamics hypothesis generation and utilization following retrieval from long-term memory.

Acknowledgments

This research was supported by a grant from the US National Science Foundation (#SES-1024650) awarded to Rick Thomas, Nicholas Lange, and Eddy Davelaar.

References

Anderson, N. H. (1973). Serial position curves in impression formation. *Journal of Experimental Psychology, 97*(1), 8-12.

Davelaar (2007). Sequential retrieval and inhibition of parallel (re)activated representations: A neurocomputational comparison of competitive cueing and resampling models. *Adaptive Behavior (15)*1, 51-71.

Davelaar, E. J., Goshen-Gottstein,Y, Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112*(1), 3-42.

Dougherty, M.R.P., Gettys, C. F., & Ogden, E. E. (1999). A memory processes model for judgments of likelihood. *Psychological Review, 106*(1), 180-209.

Dougherty, M. R. P., & Hunter, J. E. (2003)a. Probability judgment and subadditivity: The role of WMC and constraining retrieval. *Memory & Cognition, 31*(6), 968-982.

Dougherty, M. R. P., & Hunter, J. E. (2003)b. Hypothesis generation, probability judgment, and working memory capacity. *Acta Psychologica, 113*(3), 263-282.

Dougherty, M. R., Thomas, R. P., & Lange, N. (2010). Toward an integrative theory of hypothesis generation, probability judgment, and information search. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 52). San Diego: Elsevier.

Hintzman, D. L. (1986). "Schema Abstraction" in a multiple-trace memory model. *Psychological Review, 93*,411-428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528-551.

Mehlhorn, K., Taatgen, N. A., Lebiere, C., & Krems, J. F. (2011). Memory activation and the availability of explanations in sequential diagnostic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1391-1411.

Sprenger, A. & Dougherty M.R. (2011). Generating and evaluating options for decision making: The impact of sequential presented evidence. *Journal of Experimental Psychology: LMC* doi: 10.1037/a0026036

Sprenger, A. & Dougherty, M. R. (2006). Differences between probability and frequency judgments: The role of individual differences in working memory capacity. *Organizational Behavior and Human Decision Processes, 99*(2), 202-211.

Sprenger, A. M., Dougherty, M. R., Atkins, S. M., Franco-Watkins, A. M., Thomas, R. P., Lange, N., & Abbs, B. (2011). Implications of cognitive load for hypothesis generation and probability judgment. *Frontiers in Psychology-Cognitive Science, 2*, 1-15

Thomas,R. P, Dougherty, M. R., Sprenger, A. M., & Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review, 115*(1), 155-185.