Working memory dynamics bias the generation of beliefs: The influence of data presentation rate on hypothesis generation

Nicholas D. Lange · Rick P. Thomas ·
Daniel R. Buttaccio · David A. Illingworth ·
Eddy J. Davelaar

Published online: 28 September 2012
© Psychonomic Society, Inc. 2012

Abstract Although temporal dynamics are inherent aspects of diagnostic tasks, few studies have investigated how various aspects of time course influence hypothesis generation. An experiment is reported that demonstrates that working memory dynamics operating during serial data acquisition bias hypothesis generation. The presentation rate (and order) of a sequence of serially presented symptoms was manipulated to be either fast (180 ms per symptom) or slow (1,500 ms per symptom) in a simulated medical diagnosis task. When the presentation rate was slow, participants chose the disease hypothesis consistent with the symptoms appearing later in the sequence. When the presentation rate was fast, however, participants chose the disease hypothesis consistent with the symptoms appearing earlier in the sequence, therefore representing a novel primacy effect. We predicted and account for this effect through competitive working memory dynamics governing information acquisition and the contribution of maintained information to the retrieval of hypotheses from long-term memory.

Keywords Working memory · Decision making ·
Cued recall · Temporal dynamics · Hypothesis generation

In most domains, decision makers are required to generate hypotheses to explain patterns of data that have been

acquired over time. Since temporal dynamics are inherent in hypothesis generation tasks, it is important to understand how aspects of time course influence the generation of beliefs. Take as an example the plight of a medical diagnostician who must maintain multiple patient symptoms (i.e., data) observed over time when generating disease hypotheses.

Recent research concerning such temporal dynamics has investigated how different sequences of information presentation influence hypothesis generation and probability judgments. Kwan, Wojcik, Miron-shatz, Vortuba, and Olivola (2012) showed that streaks of diagnostic symptoms on a symptom–disease checklist led to higher risk estimates of having a particular disease than when the diagnostic symptoms were more uniformly distributed across the checklist. More germane to the present investigation is the finding from Sprenger and Dougherty (2012) that data (i.e., cues) occurring at the end of a serial presentation exerted more influence on hypothesis generation than did cues occurring earlier in the sequence (see also Lange, Thomas, & Davelaar, 2012). The explanation for this “recency” effect was that the last few pieces of data were more readily available in working memory (WM) and received more weight when hypotheses were retrieved (i.e., generated) from long-term memory (LTM). Relatedly, Mehlhorn, Taatgen, Lebiere, and Krems (2011) examined how the memory activation possessed by candidate hypotheses is influenced as information is acquired over time.

Our research addresses how WM dynamics govern the maintenance and contribution of data to hypothesis generation processes as data are acquired sequentially. Our theoretical framework, HyGene (Dougherty, Thomas, & Lange, 2010; Thomas, Dougherty, Sprenger, & Harbison, 2008), specifies the cognitive processes underlying hypothesis generation. Importantly, HyGene assumes that hypothesis generation represents a generalized case of cued recall whereby

N. D. Lange (✉) · E. J. Davelaar
Department of Psychological Sciences, Birkbeck College,
University of London,
Malet Street,
London WC1E 7HX, UK
e-mail: ndlange@gmail.com

R. P. Thomas · D. R. Buttaccio · D. A. Illingworth
Department of Psychology, University of Oklahoma,
Norman, OK, USA

information observed in the environment (e.g., patient symptoms) serves as cues that a decision maker uses to retrieve hypotheses from LTM.

Although HyGene has garnered some success in illuminating the underlying role of hypothesis generation in probability judgment and hypothesis testing (Dougherty et al., 2010; Thomas et al., 2008), the theory as currently formulated fails to consider how WM dynamics may influence the contributions of observed information to hypothesis generation. Moreover, the processes governing the contribution of specific information to hypothesis generation processes is not yet well understood. Fortunately, there has been considerable work done addressing WM processes in list recall that may help us bridge this gap. We look to a recent model of WM dynamics, the context-activation model (CAM; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005), to elucidate how the time course of data acquisition may influence the retrieval of hypotheses from LTM.

The CAM makes fine-grained predictions concerning the rise and fall of individual item activation levels occurring during the study phase of list recall experiments. The activation level of each item determines whether it is considered in or out of WM and is governed by three factors at each time step: bottom-up input, lateral inhibition from competing items, and self-recurrent activation. As an item is presented to the model, its corresponding memory representation receives bottom-up input driving its activation upward. When more than one item has been presented to the model, the summed activations of competing items dictate the amount of lateral inhibition applied to each item. Additionally, activation recycles back into each memory representation, providing a means of maintenance in WM in the absence of bottom-up input.

Although these dynamics have been demonstrated to adequately capture WM processes occurring in list memory paradigms, it remains an open question as to whether or not these dynamics will scale up to the generation of beliefs underlying decision making. To test this, we build upon a unique and intriguing prediction of the CAM for which evidence has been found in cued and free recall (Davelaar et al., 2005; Usher, Davelaar, Haarmann, & Goshen-Gottstein, 2008). Specifically, the CAM predicts that the rate at which items are presented radically influences the activation levels of items as they are sequentially presented to the model. This relationship is displayed in Fig. 1, which depicts two trials in which the CAM has been presented five items at two different presentation rates. Each curved line in the figure represents an item's activation, which rises and falls in concert with the competitive WM dynamics of the model. The top panel demonstrates a slow presentation rate condition that results in an “early-in early-out” pattern (recency) of informational flow through WM. The bottom panel depicts a fast presentation rate condition, demonstrating an “early-in

Fig. 1 Activation trajectories for five sequentially presented items in the context activation model in which the items have been presented at a slow rate (top panel, 1500 iterations) and a fast rate (bottom panel, 100 iterations). The model assumes a working memory threshold of 0.2. Items with activations falling below 0.2 are considered to not occupy working memory. $F(x)$ = Working Memory Activation

stay-in” pattern (primacy) where items presented early in the sequence maintain greater activation levels than those presented later in the sequence.

These predictions of primacy and recency as a function of presentation rate have been verified empirically for cued recall (Davelaar et al., 2005) and free recall (Usher et al., 2008). Under free recall, a slightly more complex pattern was found under a fast presentation rate, since one-item recency was observed in addition to primacy. We utilize this theoretical prediction and empirical evidence as a basis for predicting what information will be available in WM to support hypothesis generation when information is presented quickly or slowly. Since we assume that hypothesis generation is a generalized case of cued recall, we predict that varying the presentation rate will influence the information available in WM that serves as cues for the retrieval of hypotheses. Although recency effects in hypothesis generation have been demonstrated previously (Sprenger & Dougherty, 2012), we predict that when data are presented quickly, a shift toward primacy will occur, due to the earlier data contributing more to the generation process as a result

of competitive WM dynamics governing information acquisition.

Method

Participants

Two hundred thirty-two participants from the University of Oklahoma participated for course credit.

Design and procedure

A 2×2 between-subjects design was used in which symptom presentation rate and symptom order were manipulated (58 participants per condition). The presentation rates used were 180 ms (fast) or 1,500 ms (slow) per symptom. Table 1 presents the statistical ecology defining the associations between the various symptoms resulting from five tests and the two disease states. Each test was associated with two complementary results (positive or negative). The probabilities appearing in the table represent the likelihood of the result being in a positive state given the disease. Note that we use the terms “test result” and “symptom” synonymously within this article.

The ecology specified in Table 1 was designed such that two symptoms would suggest Disease A (tests 1 and 2) and two would suggest Disease B (tests 4 and 5). The second independent variable was the symptom order presented in the elicitation. In one condition, the symptoms were presented in the order in which they appear in Table 1 ($1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5$), and in the other condition, the nondiagnostic (ND) symptom 3 was moved to the end, resulting in the ordering $1 \rightarrow 2 \rightarrow 4 \rightarrow 5 \rightarrow 3$. We refer to these orderings as ND-Middle and ND-Last. The ND-Last ordering was included in anticipation of the one-item recency observed in the fast condition of Usher et al. (2008). Although the last item may be maintained in WM under the fast rate (in addition to the initial items), its nondiagnosticity makes it a neutral cue, and it therefore should not bias generation toward one hypothesis or the other. Since the same symptoms were always presented at elicitation in every condition, the objective posterior probabilities of the hypotheses were always equal.

Table 1 Symptom-Disease ecology in which the results of tests 1 and 2 are associated with Disease A and results of tests 4 and 5 are associated with Disease B. Note that we use the terms test result and symptom interchangeably for the purposes of this experiment

	Test 1	Test 2	Test 3	Test 4	Test 5
Disease A	0.9	0.7	0.5	0.3	0.1
Disease B	0.1	0.3	0.5	0.7	0.9

Prior to the beginning of the experiment, the two fictitious disease names used in the experiment (Metalytis or Zymosis) were randomly assigned to the rows of Table 1 (i.e., Disease A or Disease B). Additionally, the test nomenclatures—for example, CUL (bacterial culture)—were randomly assigned to the columns of Table 1 (i.e., tests 1–5), as were the “positive” and “negative” result assignments within each test. So although the ecology was consistent across participants, the labels associated with the diseases, tests, and symptom states were randomized.

The experiment consisted of two phases. The first phase was an exemplar training phase in which a series of hypothetical prediagnosed patients (100 for each disease) was presented to the participants in order for them to learn, through experience, the contingencies between the diseases and symptoms displayed in Table 1. Each of these patients was represented by the disease name at the top of the screen, the presenting symptom of severe abdominal pain, and a series of simultaneously presented symptoms resulting from each of the five tests. An example of a training patient exemplar is presented in Fig. 2, in which all symptoms appeared at the same time. The five tests were described to the participants in the instructions as bacterial culture, temperature, balance, vision, and an eardrum test, corresponding to the abbreviations appearing in Fig. 2. The vertical position of each test and its resulting symptom state on the exemplars was randomized for each participant. Each exemplar was presented to the participants for 5 s, at which point they were prompted to enter the first letter of the current patient’s disease when they were ready to continue. Memory checks appeared at random, requiring a forced choice response regarding one of the previous patient’s symptoms or his or

Fig. 2 Example exemplar used in training phase

her disease state, for which immediate feedback was provided. These were included to incentivize attention in the training task. Following the exemplar training, a short diagnosis test was administered by presenting the participants with a single symptom and asking them to report the most likely disease, given the symptom.

The second phase of the experiment was the elicitation phase, in which our experimental design was implemented. The participants were instructed that they would now be diagnosing a patient who has come to them in need of diagnosis. Prior to the critical trial of interest, the participants were presented with examples of the serial symptom presentation in which numbers appeared in place of symptoms, either fast or slow. Although all participants were shown examples of presentation rates, this was done so that those in the fast presentation rate condition were not caught off guard by the speed of symptom presentation on the forthcoming critical trial. At the participants' readiness, they triggered the onset of the patient's symptoms, which were presented sequentially at either the fast or slow presentation rate and in one of the two symptom orderings. Following the presentation of the last symptom, a prompt appeared on the screen asking "Most Likely Disease?" to which the participants responded by pressing that disease's first letter (either M or Z), as they had done in the training and were instructed to do prior to the presentation of the patient's symptoms.

As was discussed above, our primary prediction was that the presentation rates of the symptoms would influence which hypothesis would be generated/selected as most likely. Specifically, we predicted that when the symptoms were presented rapidly, there should be a preference toward the disease consistent with symptoms presented early in the sequence and, conversely, that when the symptoms were presented slowly, we should observe a preference for the disease consistent with the symptoms appearing later in the sequence. We further expected that the manipulation of cue ordering would strengthen the effect of the presentation rate by removing the contribution of the last item to the generation process under the fast presentation rate for the ND-Last ordering.

Results

To assess performance in the training phase, we analyzed the data from the random memory checks and the diagnosis test directly preceding the elicitation phase. Performance on the random memory checks was very good, with an average across participants of 91 % ($SD = 0.1$) and a median of 94 % (interquartile range: 0.88–0.97). Performance on the diagnosis test was fair, with an average across participants of 68 % ($SD = 0.23$) and a median of 75 % (interquartile range: 0.5–0.88). These results indicate a high level of attention

allocated to the exemplars during the training phase and reasonable learning of the disease–symptom associations.

To test the effects of presentation rate, data order, and their potential interaction, a 2×2 logistic regression was carried out on the proportion of participants who selected Disease A. As can be seen in Fig. 3, a main effect of presentation rate was obtained, since participants were more likely to choose Disease A when the symptoms were presented rapidly (and, conversely, more likely to choose Disease B when the symptoms were presented slowly), $\chi^2(1) = 6.28$, $p < .05$, with no effect of cue order, $\chi^2(1) = 0.15$, $p = .69$. Although the effect of presentation rate looks stronger for the ND-Last order, as compared with the ND-Middle order, the interaction between presentation rate and order was nonsignificant, $\chi^2(1) = 0.86$, $p = .35$.

Discussion

We leveraged theory and empirics concerning WM dynamics in traditional memory paradigms to predict the influence of presentation rate on the generation of beliefs in a simulated medical diagnosis task. The theoretical stance forwarded here is that the same WM dynamics governing item activation in list recall tasks adequately describe the WM dynamics governing data acquisition in support of hypothesis generation tasks. In accordance with the WM dynamics of the CAM, we predicted that later items would more often occupy WM at the end of item presentation under slow rates and, therefore, exert more influence over generation in such conditions. This proposition readily accounts for the recency effects demonstrated by Sprenger and Dougherty (2012). Likewise, we observed recency in the present experiment under the slow presentation rate. Specifically, when the serial presentation of five symptoms was presented at a slow rate (1,500 ms per symptom), the symptoms appearing later in the sequence exerted more influence on diagnosis than did the symptoms appearing earlier in the sequence. Crucially, in addition to this finding, we demonstrated a theoretically and empirically novel effect whereby primacy was observed in which symptoms appearing earlier in the sequence exerted more influence on diagnosis under the fast presentation rate (180 ms per symptom).

The experiment clearly demonstrates that whether primacy or recency obtains in a hypothesis generation task depends on dynamic WM maintenance processes that are sensitive to the temporal characteristics of the task. Since neither the presentation rate of the data nor the ordering of the data influence the posterior probabilities according to normative theory, these effects represent temporal biases in hypothesis generation. Our theoretical position that only the information being actively maintained in WM contributes to hypothesis generation is consistent with

Fig. 3 Proportion of generation of Disease A vs. Disease B by presentation rate and symptom ordering

the findings of the present experiment. Thus, merging work in hypothesis generation with work investigating the WM dynamics in list recall provides a straightforward theoretical account of the primacy and recency effects demonstrated in our simulated medical diagnosis task.

It is potentially important to note, however, that one possible prediction stemming from the combination of the CAM and the empirically observed one-item recency effect is that the recency effect under the slow presentation rate would be stronger under the ND-Middle ordering, in comparison with the ND-Last ordering. The present data do not provide evidence for this, since the recency effects are rather equivalent across the cue ordering conditions. Whereas the CAM predicts that decision makers will primarily utilize the last two pieces of observed data (i.e., symptoms) under slow presentation rates, the results from the present experiment suggest that people may most often use the last three pieces. This interpretation readily accounts for the similarity in the recency effects across order conditions (under slow presentation rates), while allowing for the observed divergence in primacy (under fast presentation rates).

Since temporal constraints are inherent in our deployment of hypothesis generation in real-world tasks, the bias demonstrated in the present experiment highlights the importance of understanding how internal memory dynamics and external temporal task characteristics interact to influence hypothesis generation processes. Although we relied upon a merger between a model of WM dynamics (Davelaar et al., 2005) and a model of hypothesis generation (Thomas et al., 2008) to predict how the presentation rate of evidence would influence hypothesis generation, it is interesting to consider to what extent the present data can be interpreted through the lens of sequential-sampling models of choice. Consistent with many choice tasks, the particular simulated medical diagnosis task employed here provided only two possible diseases and required only a single response. The sensitivity to presentation rate in the CAM is due to the balance of competition between items and the self-recurrency of each item. Given the strong mathematical similarity between the dynamic buffer operations in the

CAM (Davelaar, et al., 2005) and those in the leaky competing accumulator model (LCA; Usher & McClelland, 2001), we additionally expect that a recency-to-primacy shift with increasing presentation rate should occur across a wide range of choice response tasks.¹ We note that sequential-sampling models contrasted in the literature (e.g., Ratcliff, 1978; Usher & McClelland, 2001) do not require retrieval from memory, as we have assumed in our current theoretical position. Although sequential-sampling models have been applied to tasks in which the order of items in a sequence is manipulated (Pietsch & Vickers, 1997), the influence of presentation rate has not yet been investigated. The same is true of the belief adjustment model (Hogarth & Einhorn, 1992), which accounts for differential weighting of early versus late evidence contingent on various task characteristics but does not address evidence presentation rate. Additionally, it could be interesting to apply decision field theory (Bussemeyer & Townsend, 1993) to varying conditions of evidence presentation rate and determine what adaptations this framework would require to account for the primacy effect under the fast presentation rate observed in the present experiment. In light of the similarity between the currently employed task and existing choice tasks, the present effect of evidence presentation rate may provide valuable data for the further development and testing of dynamic models of choice, belief adjustment, and hypothesis generation.

Of additional interest for future research will be how presentation rates of greater ecological validity influence hypothesis generation. More naturalistic presentation rates (i.e., differences in seconds or minutes, rather than milliseconds) will likely lead to gradations in the contribution of recent items, as opposed to the recency–primacy shift observed here. Such differences in recency biases under more ecological rates are important to understand and will have direct application to professional domains (e.g., medical

¹ In fact, Usher and Davelaar (2002; see also Davelaar, 2007) showed how modulating the alpha (self-recurrency) and beta (global inhibition) parameters can be used to push the model into a selection mode.

diagnosis, development of decision support tools). Additionally, within our theory, the present effect should extend to hypothesis judgment and testing tasks. Such experiments will be informative for understanding the further impacts of temporal biases underlying hypothesis generation on judgment and decision making.

Author note This work was supported by NSF Grant # SES-1024650 awarded to R.P. Thomas, N.D. Lange, & E.J. Davelaar.

References

- Busemeyer, J. R., & Townsend, J. T. (1993). Decision Field Theory: A dynamic cognition approach to decision making. *Psychological Review*, 100, 432–459.
- 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, 3–42.
- Dougherty, M., Thomas, R., & Lange, N. (2010). Toward an integrative theory of hypothesis generation, probability judgment, and hypothesis testing. In B. H. Ross & B. H. Ross (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol 52) (pp. 299–342). San Diego, CA US: Elsevier Academic Press.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief- adjustment model. *Cognitive Psychology*, 24(1), 1–55.
- Kwan, V. S. Y., Wojcik, S. P., Miron-shatz, T., Vortuba, A. M., & Olivola, C. Y. (2012). Effect of Symptom Presentation Order on Perceived Risk. *Psychological Science*. doi:10.1177/0956797611432177
- Lange, N. D., Thomas, R. P., & Davelaar, E. J. (2012). Temporal dynamics of hypothesis generation: Influences of data serial order, data consistency, and elicitation timing. *Frontiers in Psychology: Cognitive Science*, 3, 215. doi:10.3389/fpsyg.2012.00215
- 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, 1391–1411.
- Pietsch, A., & Vickers, D. (1997). Memory capacity and intelligence: Novel techniques for evaluating rival models of a fundamental information-processing mechanism. *The Journal of General Psychology*, 124, 229–339.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Sprenger, A., & Dougherty, M. P. (2012). Generating and evaluating options for decision making: The impact of sequentially presented evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 550–575.
- Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review*, 115, 155–185.
- Usher, M., & Davelaar, E. J. (2002). Neuromodulation of decision and response selection. *Neural Networks*, 15, 635–645.
- Usher, M., Davelaar, E. J., Haarmann, H., & Goshen-Gottstein, Y. (2008). Short Term Memory After All: Comment on Sederberg, Howard, and Kahana (2008). *Psychological Review*, 115, 1108–1118.
- Usher, M., & McClelland, J. L. (2001). The Time Course of Perceptual Choice: The Leaky Competing Accumulator Model. *Psychological Review*, 108, 550–592.