Event Cognition and Holistic versus Fragmented Remembering and Forgetting

Daniela Parra (dparra2@nd.edu)
Department of Psychology, University of Notre Dame
Notre Dame, IN 46656 USA

Nicole Antes (n.antes@iwm-tuebingen.de)
Leibniz-Institut für Wissensmedien
Tübingen, Germany

Gabriel A. Radvansky (gradvans@nd.edu)
Department of Psychology, University of Notre Dame
Notre Dame, IN 46656 USA

Abstract

This study assessed the holistic and fragmented retention and forgetting of event models. We report four experiments that manipulated causality, co-reference, events versus objects, and description determinacy. While increased causal connections among events increased holistic remembering, there was no clear effect for manipulations of co-reference, events versus objects, or determinacy. Thus, our work suggests that there are limits to the extent to which different types of events are remembered and forgotten in a holistic or fragmented manner. That said, all of our event did show significantly greater than chance holistic remembering, suggesting that the very act of creating event models leads these memories to be remembered or forgotten as wholes to a greater extent.

Keywords: event cognition; memory; holistic remembering; forgetting

Introduction

People forget things. All the time. The more time that passes, the worse it gets (Ebbinghaus, 1885). However, not everything is remembered or forgotten in the same way. For example, Fisher and Radvansky (2018) had people read stories and then tested memory immediately and up to 12 weeks later. Verbatim memories were forgotten soon after reading. Memories for the propositional textbase, the abstract ideas in the texts, were forgotten more slowly. After a week, there was a large drop in memory, and this type of memory grew worse. Finally, event model memories, which reflect understanding of what a text was about, and include both information in the texts and inferences drawn from general world knowledge, were retained with little forgetting over the 12-week period. Thus, remembering and forgetting can be very different for different types of knowledge.

Our focus is on the event model level (Johnson-Laird, 1983; Radvansky & Zacks, 2014; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Rather than representing the individual components of experience (e.g., sentences of a text), event models are more holistic, analog, and embodied. Each model can serve as a unit. This has implications for retention and forgetting. As suggested by Horner & Burgess (2013), event models are more likely to be remembered and forgotten as wholes. People would either remember an event in its entirety, or forget it completely.

The approach, across many studies (e.g., Bisby, Horner, Bush, & Burgess, 2018; Horner, Bisby, Bush, Lin, & Burgess, 2015; Horner & Burgess, 2013; 2014; Joensen, Gaskell, & Horner, 2020), was to give sets of items, often person-object-location triples, and test memory. Memory was tested using six alternative recognition tests. Often, one concept was given as a cue, such as the person from a triple, and six alternatives from the same category would then be presented with it, such as six locations. One alternative was studied with the cue and the others were not. This was done for all six pairs for each triple (e.g., person-object, person-location, object-person, object-location, location-person, location-object). From this, Horner developed an analysis protocol that uses recognition data to quantify the degree to which events were remembered holistically relative to chance (Bisby, et al., 2018; Horner & Burgess, 2013). This is called **dependency**, and the outcome of this protocol is a **dependency score**. Assuming complete dependency, people would remember either all of the targets associated with all three cue types from an event, with all elements being dependent on the others, or none of them, with correct responses within a type only occurring at chance levels.

This idea of more holistic event model memory has been tested by varying the experience of the original information (Horner & Burgess, 2014). For each triple, the items were presented in either a closed loop (all three pairs were given), or an *open loop* (only two of the three were). Dependency is greater for closed than open loop conditions. That is, when people experience all of an event, it is more likely to be integrated, and remembered or forgotten as a whole.

The aim of our project was to further explore the idea that event models are more likely to be remembered and forgotten as wholes. We went beyond the open-/closed-loop approach by looking at other factors involved in the creation of integrated event models. The prediction in all cases is that increased integration leads to greater memory dependency, suggesting that event models are more likely to be remembered or forgotten as wholes.

In pursuit of this aim, we did four experiments that manipulated causality, co-reference, events versus objects,

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

and determinacy. Experiment 1 assessed memories for previously-read real-world novels. The prediction was that events that are more causally connected with others in the novel will be more integrated, and so, will have greater dependency. Experiment 2 compared memory for statements that could be integrated into a common event model (i.e., two objects in one location) with those that could not (i.e., one object in two locations). The prediction was that greater integration would result in greater dependency. Experiment 3 was based on a claim by Andermane, Joensen, and Horner (2021) that event memories would show greater dependency than memories for objects. Experiment 4 tested memories for spatial descriptions that were either determinate (a description fits one spatial layout) or indeterminate (it fits two layouts). The prediction was that dependency would be greater for determinate than indeterminate descriptions. Overall, our project assessed the holistic remembering and forgetting of event memories using multiple methods and measures to converge on this issue.

Experiment 1

Experiment 1 assessed dependency for event models created from reading real-world novels. Events with more causal relevance are remembered better (Doolen & Radvansky, 2021). The prediction here was that events that were better integrated into the novel, as defined by the level of causal connectivity, would show greater retrieval dependency.

Method

Participants We tested 85 college students (51 female), 18-27 years old (M = 19.2; SE = .16). Informed consent was obtained and our Institutional Review Board approved all procedures. People who reported having read a selected novel more than once were excluded from analyses (4 people). There were 24 participants who had their memory probed for *1984*, 18 for *Lord of the Flies*, 23 for *The Scarlet Letter*, and 31 for *To Kill a Mockingbird*.

Materials The following books were chosen because they are commonly read in American high schools: *1984* (Orwell, 1949), *Lord of the Flies* (Golding, 1954), *The Scarlet Letter* (Hawthorne, 1850), and *To Kill a Mockingbird* (Lee, 1960). Prior work (Doolen & Radvansky, 2021) identified 22 events for each and calculated the causal connectivity among them. Events with three or fewer causal connections were classified here as low connectivity (M = 1.9; SE = .13), and those with four or more were classified as high connectivity (M = 7.4; SE = 0.9). We derived probe triples for 22 events in each novel. Each triple was made of the best three concepts from a set of people, objects, locations, and activities. Care was taken to ensure triple elements across events did not overlap.

Each probe was composed of one concept from a triple as a cue, along with six options from a category. For example, if the cue was a location, all of the options could be objects (correct + 5 distractors). As instance, for *1984*, a location cue "apartment" could be paired with a target object, "diary", and five other distractor objects, such as **"**Julia's picture", "glass paperweight", "10th Edition of Newspeak Dictionary", "Brotherhood", and "note" that were also in the book, but not in the apartment. Each cue type was paired with each of its two target types, for a total of six probes per triple.

Procedure and Analysis People first indicated which of the 4 novels they had read. Those who had not read any were dismissed. The rest were then asked if they had actually read the entire novel and if so, whether they had watched a film or a theater production of it. From the novels a person actually read without the influence of a film or theater production, one was randomly selected. Participants then did a six-alternative recognition test made up of 132 memory probes. Both the cues and the response choices were presented in a random order for each participant. People also reported the number of times they had read the novel and how long it had been since the last reading of it (in months and years).

The data were analyzed in two ways. First, we analyzed accuracy in terms of delay and causal connectivity. Next, and of particular interest here, we calculated the Horner and Burgess (2013) dependency measure. This analysis also produces an independent model, which acts as an index of chance dependency, taking into account the level of accuracy. The dependency measure is contrasted with the independent model to assess whether performance is beyond that expected by chance. Details of the mathematical formulae of the protocol are provided by Horner and Burgess.

Results

Overall accuracy was .35 (SE = .01). There was a wide range of delays between reading and testing, anywhere from 0.2 to 13.1 years (M = 4.5; SE = .2). Although a standard finding in memory research is a decline with longer retention intervals, that was not observed here, as can be seen in Figure 1. There was no correlation between retention interval and accuracy, r = -.14, p = .19. This lack of change over time is consistent with other findings in the literature in which there is little to no change in overall memory performance over the course of many years (e.g., Bahrick, Bahrick, & Wittlinger, 1975; Conway, Cohen, & Stanhope, 1991; Haist, Gore, & Mao, 2001). As such, we do not consider delay further.

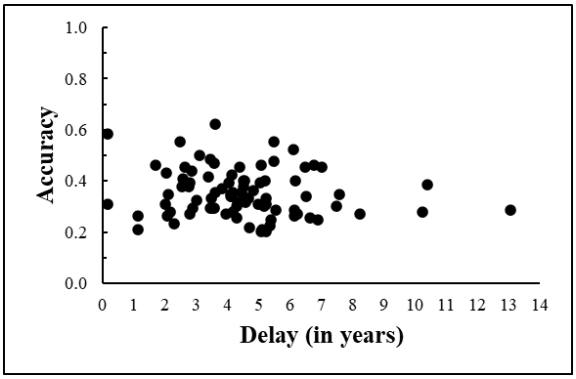

Figure 1: Accuracy across retention intervals

We assessed accuracy as a function of high and low causal connectivity. This was significant, $F(1,84) = 14.21$, $p < .001$, $\eta_p^2 = .15$, with greater accuracy for the high ($M = .42$; $SE = .02$) than the low connectivity probes ($M = .34$; $SE = .01$), consistent with prior work (Doolen & Radvansky, 2021).

For the Horner and Burgess (2013) dependency score, there was also no relationship with delay, $r = .08$, $p = .65$ (see Figure 2). In other words, the level of dependency, of holistic retrieval, for these events remained constant over the years. Because of this, delay is not considered further.

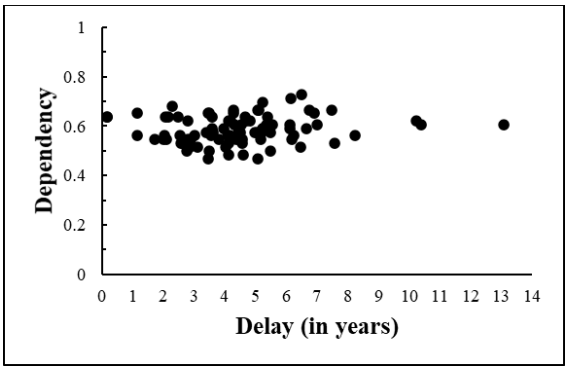

Figure 2: Dependency across retention intervals

Dependency results are shown in Figure 3. The scores were submitted to a 2 (Dependency: Data, Independent Model) X 2 (Causal Connectivity: Low, High) repeated measures ANOVA. There was a main effect of Dependency, $F(1,84) = 33.46$, $p < .001$, $\eta_p^2 = .29$, with novel events being remembered more holistically than chance. The effect of Causal Connectivity was not significant, $F < 1$, but the interaction was, $F(1,84) = 6.49$, $p = .01$, $\eta_p^2 = .07$. Simple effects tests showed Dependency for both low, $F(1,84) = 8.29$, $p = .005$, $\eta_p^2 = .09$, and high connected events, $F(1,84) = 25.85$, $p < .001$, $\eta_p^2 = .24$, although more so for high events. Thus, the importance of an event was related to it being more likely to be stored in a holistic manner. In summary, the results of Experiment 1 were consistent with the prediction that more integrated event models would show more evidence of holistic, rather than fragmented, forgetting.

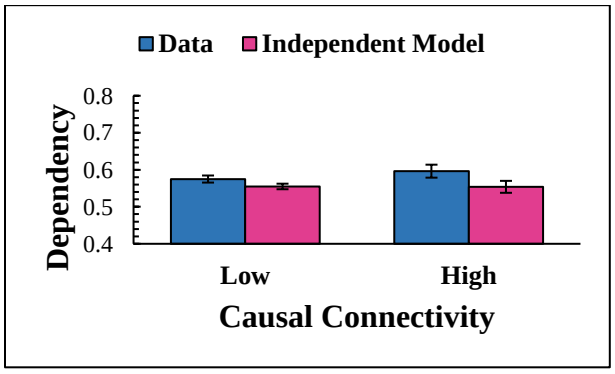

Figure 3: Dependency for Low and High connected events

Experiment 2

Experiment 2 explored dependency when people learned sentences that were likely to be integrated into an event model or not. Moreover, there was no explicit instruction to imagine items interacting. This is in line with work on the differential fan effect which has shown model creation, integration, and use during retrieval for sentences, even in the absence of explicit instructions to do so (e.g., Radvansky, et al., 2017; Radvansky & Zacks, 1991). The prediction was that dependency would be greater in the Single Location condition, in which information integrated into a single model is more likely, than the Multiple Location condition, in which separate information model storage is more likely.

Method

Participants We tested 123 college students (80 female), 18-27 years old ($M = 19.2$; $SE = .11$). Informed consent was obtained and our Institutional Review Board approved all procedures. The data from 6 additional people was dropped because of chance (.167) or below accuracy.

Materials The study sentences were of the form "*The object is in the location.*" Each list was generated by randomly assigning objects and locations into a triple. For the Single Location condition, there were two objects associated with one location. For example, two sentences in this condition could be "The welcome mat is in the museum." and "The potted palm is in the museum." For the Multiple Location condition, there was one object associated with two locations. For example, two sentences in this condition could be "The waste basket is in the airport." and "The waste basket is in the hotel." There were four triples for each condition and each triple was shown twice during study. As in Experiment 1, six memory probes were created for each triple for a total of 48. For each probe, the five distractors were from the same category as the target (either objects or locations).

Procedure People were first given a list of 16 sentences to study. These were presented one at a time, for 7 seconds each. Following this, they were given a recognition test in which a cue was presented with six response alternatives. They were to indicate which alternative was associated with the cue. Probe and response orders were randomized for each person.

Results

To assess accuracy in the Single and Multiple Location conditions, the data were submitted to a repeated measures ANOVA. This difference was significant, $F(1,122) = 5.14$, $p = .03$, $\eta_p^2 = .04$, with people being more accurate for the Single Location condition ($M = 0.52$, $SE = 0.02$) than the Multiple Location condition ($M = 0.48$, $SE = 0.02$). This is consistent with prior work (e.g., Radvansky & Zacks, 1991). This is because information can be integrated into a single mental model, thereby mitigating sources of retrieval interference.

The results of the dependency analysis are shown in Figure 4. These data were submitted to a 2 (Dependency: Dependent, Independent Model) X 2 (Model: Single vs. Multiple Location) repeated measures ANOVA. There was a main effect of Dependency, $F(1,122) = 19.36$, $p < .001$, $\eta_p^2 = .14$, with scores being higher than the independent model. There was also a main effect of Model, $F(1,122) = 4.90$, $p = .03$, $\eta_p^2 = .04$, with the scores being higher for the Single Location than Multiple Location condition, consistent with the idea that there was more integration in the Single Location condition. However, the interaction was not significant, $F(1,87) = 0.91$, $p = .34$, $\eta_p^2 < .01$, with a similar integration benefit across the two conditions. Although the dependency scores were higher in the Single Location condition, the benefit differed from chance to the same degree as in the Multiple Location condition. Thus, we did not find conclusive support that there was greater holistic forgetting in the Single Location condition, contrary to the prediction.

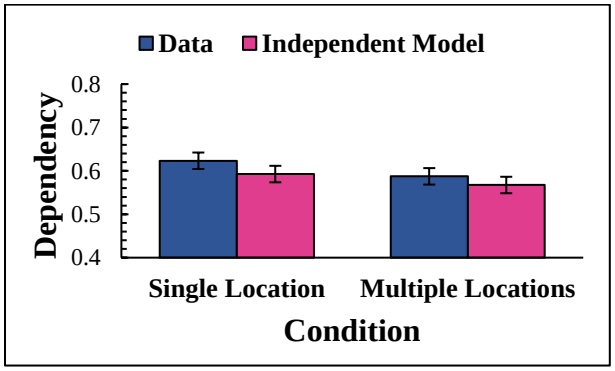

Figure 4: Recognition test dependency scores across Single and Multiple Location Conditions

Experiment 3

Experiment 3 assessed whether the integration of concepts that make up an event is stronger than for concepts that make up an object. Andermane et al. (2021) suggest that object memories are more likely to be remembered in a fragmented way than event memories because events and objects are processed by different brain regions. Their argument was that events are more likely to be processed by the hippocampus, which binds together different event elements, whereas objects are more likely to be processed by the perirhinal cortex, which involves object features being encoded independently. The prediction is that event memories will show greater dependency that object memories.

Method

Participants We tested 47 students (25 female), 18-23 years old ($M = 19.1$; $SE = .09$). Informed consent was obtained and our Institutional Review Board approved all procedures.

Materials *Event triples* were made up of person-object-location concepts, as has been done previously (e.g., Horner & Burgess, 2013). An example of an event triple could be Oprah Winfrey – stroller – castle. For *object triples*, two adjectives and one object were used. One adjective was a physical property (e.g., old, ugly, green), while the other was an abstract property (e.g., purchased, rare, cheap). An example of an event triple could be comb – gray – eco-friendly. Concepts for each triple were randomly selected for a total of 60 triples. Memory probes for the six-alternative recognition test were created in a similar way as was done in Experiments 1 and 2.

Procedure Participants were asked to imagine, as vividly as possible, each group of concepts interacting in a meaningful way. The event and object triples were presented in a random order for 9 seconds each. After study, people were given a 360 item, six-alternative recognition test.

Results

The accuracy data were submitted to a repeated measures ANOVA. This difference was significant, $F(1,46) = 54.57$ $p < .001$, $\eta_p^2 = .54$, with people being more accurate for the Event condition ($M = 0.65$, $SE = 0.04$) than the Object condition ($M = 0.50$, $SE = 0.04$).

The results of the dependency analysis are shown in Figure 5. These data were submitted to a 2 (Dependency: Dependent, Independent Model) X 2 (Item Type: Event, Object) repeated measures ANOVA. There was a main effect of Dependency, $F(1,46) = 106.31$, $p < .001$, $\eta_p^2 = .70$, with the dependency being higher than in the independent model. There was also a main effect of Item Type, $F(1,46) = 15.90$, $p < .001$, $\eta_p^2 = .26$, with the scores being higher for events than objects. The interaction was not significant, $F(1,46) = 1.56$, $p = .22$, $\eta_p^2 = .03$. This suggests that the integration benefit was similar across the two conditions. This was true even if one considers that the object descriptions may have been odder than the event descriptions. So, like Experiment 2, we did not find conclusive support that there was greater holistic forgetting in one condition over the other, contrary to the prediction.

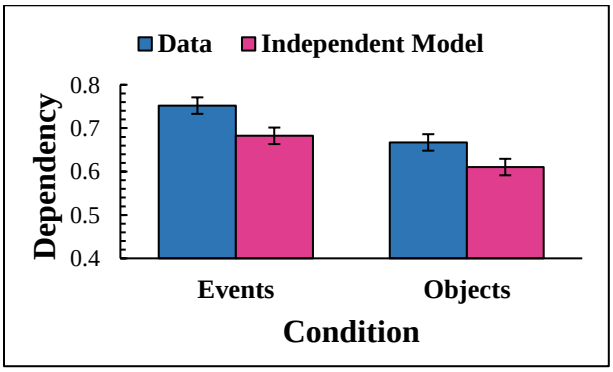

Figure 5: Recognition test dependency across Event and Object triples.

Experiment 4

Experiment 4 aimed to assess whether different types of spatial descriptions lead to different levels of dependency. Prior work looked at how people remember spatial information when presented as indeterminate or determinate descriptions (Mani & Johnson-Laird, 1982). *Determinate descriptions* are consistent with one spatial layout, and lead to more use of event models. In contrast, *indeterminate descriptions* are consistent with multiple layouts and lead to more use of propositional textbase knowledge. We expected greater dependency for determinate descriptions.

Method

Participants We tested 55 students (41 female), 18-22 years old (M = 19.0; SE = .13). Informed consent was obtained and our Institutional Review Board approved all procedures.

Materials The spatial descriptions were similar to those used by Mani and Johnson-Laird (1982). Each consisted of three sentences with 4 concepts: person, object, animal, and plant. Each sentence conveyed the relation between two items. We had 8 determinate and 8 indeterminate descriptions. For **determinate descriptions**, the spatial layouts were unambiguous. Thus, these descriptions were consistent with a single event, and would be represented by a single event model. An example determinate description is as follows, along with a corresponding layout:

The A is behind the B.	A D C	The pumpkin is behind the computer.
The A is to the left of the D.	B	The pumpkin is to the left of the elephant.
The C is to the right of the D.		The lawyer is to the right of the elephant.

Indeterminate descriptions corresponded to two layouts. Thus, these descriptions were consistent with multiple events, and would not be effectively captured by a single event model. Thus, people would be more likely to be dependent on propositional memories. For example:

The A is behind the B.	A D C	or	A C D	The pumpkin is behind the computer.
The A is to the left of the D.	B		B	The pumpkin is to the left of the elephant.
The C is to the right of the A.				The lawyer is to the right of the pumpkin.

Person, object, animal, and plant concepts were randomly assigned from a pool of 16 each. We combined these into 16 sets. We generated 12 test sets for the recognition task for each description (192 in total). In these sets, every concept served as a cue, and each of the other concepts as a target. Again, distractors were from the same category.

Procedure People were asked to imagine the described spatial layouts and read the sentences at their own pace. Each description was presented sentence by sentence. After reading all of the descriptions, a short distractor task was presented followed by the six-alternative forced-choice recognition test. This test presented a cue concept along with six response options. The order of all test sets in the recognition task was randomized.

Results

The accuracy data were submitted to a repeated measures ANOVA. There was a difference between determinate and indeterminate conditions, $F(1,54) = 6.72$, $p = .01$, $\eta_p^2 = .01$. People had higher accuracy for indeterminate (M = 0.26, SE = 0.02) than determinate (M = 0.23, SE = 0.01) descriptions. This is the opposite of what was expected.

The data for the dependency analysis were submitted to a 2 (Dependency: Dependent, Independent Model) X 2 (Description: Determinate, Indeterminate) repeated measures ANOVA (Figure 6). There was a main effect of Description with greater dependency for the indeterminate (M = 0.52, SE = 0.14) than the determinate condition (M = 0.46, SE = 0.12), $F(1,54) = 7.62$, $p = .007$, $\eta_p^2 = .03$, in line with the accuracy data. Additionally, there was a main effect of Dependency, $F(1,54) = 5.18$, $p = .02$, $\eta_p^2 = .003$, with the likelihood of memories being remembered in a more holistic manner being higher than would be expected by chance. Like Experiments 2 and 3, there was no interaction, $F(1,54) = 2.73$, $p = .10$, $\eta_p^2 = .001$. That is, whether a description was determinate or indeterminate had no impact on later memory dependency. Overall, this contradicts the prediction that the determinate description benefits from a clear spatial layout for greater dependency.

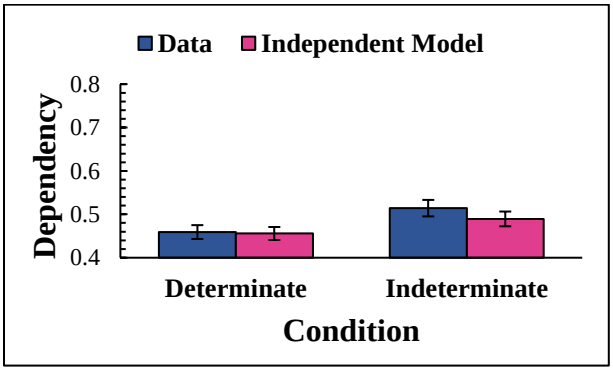

Figure 6: Recognition test dependency across determinate and indeterminate descriptions

General Discussion

Our focus was to explore the idea that integrating information into event models leads to more holistic, as compared to fragmented, remembering and forgetting. This builds off of a long line of work (Bisby et al., 2018; Grande et al., 2019; Horner et al., 2015; Horner & Burgess, 2013; 2014; James et al., 2020; Joensen et al., 2020; Ngo et al., 2019). This is important because holistic forgetting would provide some insight into qualitatively different types of memory representations, and the consequences this has for later retrieval. Thus, we were motivated to extend our theoretical understanding beyond materials such as concept triplets.

For our study, we used a multi-method approach that converged on the idea that increased event model integration would lead to greater retrieval dependency. Our studies used real-world novels, lists of sentences, event versus object descriptions, and spatial descriptions. We found support for our predictions only with novels in Experiment 1.

Why did this happen? At this point, we can only offer some suggestions. Our working hypothesis is that the primary difference between novels and other material types was the richness of the events and the role of causality, which are prevalent in novels. We saw that dependency varied with causal connectivity. All of the other materials had a more impoverished, list-like quality; lists of sentences (Experiment 2), lists of events and objects (Experiment 3), and lists of spatial layouts of random collections of items (Experiment 4). Although prior work (e.g., Horner & Burgess, 2013) used relatively impoverished sets of person-location-object triples, the experimental comparisons typically involved the completeness of the descriptions, not differences in the nature of the described events. It may be that differences in dependency for our manipulations would only be observed with either more meaningful processing of the materials (e.g., comparing descriptions with actual layouts), or with increased exposure to the materials, which makes it more likely that inferences and elaboration would occur, making it more likely that our manipulations would have an impact.

It should also be noted that in all cases, dependency scores were always greater than what would be expected by chance. Thus, people were creating and using event models to represent this information, and these event models resulted in greater holistic forgetting. In that context, there were no differences observed in the type of event models created. The use of whatever the nature of the models involved led to better-than-expected dependency overall.

In summary, there is some evidence that there are differences in the degree to which different types of memories are retrieved in a more holistic way. However, our work places limitations on the generality of this idea. There are some cases in which, although there may be variation in the degree of integration, or the adequacy, of the created event models, there is no variation in the degree to which these memories result in more holistic or fragmented forgetting. Overall, holistic remembering is more likely to be observed with more complex event memories in which there is a narrative or causal structure, and less so with more impoverished lists of materials.

References

Andermane, N., Joensen, B. H., & Horner, A. J. (2021). Forgetting across a hierarchy of episodic representations. *Current Opinion in Neurobiology, 67*, 50-57.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1), 54-75.

Bisby, J. A., Horner, A. J., Bush, D., & Burgess, N. (2018). Negative emotional content disrupts the coherence of episodic memories. *Journal of Experimental Psychology: General, 147*(2), 243–256.

Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long-term retention of knowledge acquired through formal education. *Journal of Experimental Psychology: General, 120*(4), 395-409.

Doolen, A. C., & Radvansky, G. A. (2021). A novel study: long-lasting event memory. *Memory, 29*(8), 963-982.

Ebbinghaus, H. (1885/1913). *Memory: A Contribution to Experimental Psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Columbia University, Teacher's College. (Reprinted 1964, New York: Dover).

Fisher, J. S., & Radvansky, G. A. (2018). Patterns of forgetting. *Journal of Memory and Language, 102*, 130-141.

Golding, W. (1954). *Lord of the Flies*. Faber and Faber.

Haist, F., Gore, J. B., & Mao, H. (2001). Consolidation of human memory over decades revealed by functional magnetic resonance imaging. *Nature Neuroscience, 4*(11), 1139-1145.

Hawthorne, N. (1850). *The Scarlet Letter.* Ticknor, Reed & Fields.

Horner, A. J., Bisby, J. A., Bush, D., Lin, W. J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications, 6*(1), 7462.

Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General, 142*(4), 1370–1383.

Horner, A. J., & Burgess, N. (2014). Pattern completion in multielement event engrams. *Current Biology, 24*(9), 988-992.

Joensen, B. H., Gaskell, M. G., & Horner, A. J. (2020). United we fall: All-or-none forgetting of complex episodic events. *Journal of Experimental Psychology: General, 149*(2), 230–248.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Harvard University Press.

Lee, H. (1960). *To Kill a Mockingbird*. J. B. Lippincott & Co.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition*, *10*(2), 181–187.

Orwell, G. (1949), *1984*. Secker & Warburg.

Radvansky, G. A., O'Rear, A. E., & Fisher, J. S. (2017). Event models and the fan effect. *Memory & Cognition, 45*, 1028-1044.

Radvansky, G. A., & Zacks, J. M. (2014). *Event Cognition.* Oxford University Press.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 940–953.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of Discourse Comprehension. New York: Academic Press.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162–185.