On quantifying schematicity of future narratives

Isaac Kinley (ikinley@research.baycrest.org)[1]

Donna Rose Addis (draddis@research.baycrest.org)[1,2]

Reece P Roberts (r.roberts@auckland.ac.nz)[3,4]

Samuel Fynes-Clinton (research.sfynesclinton@gmail.com)[1]

Yang Xu (yangxu@cs.toronto.edu)[5]

[1]Rotman Research Institute, Baycrest Academy of Research and Education
[2]Department of Psychology, University of Toronto
[3]School of Psychology, Faculty of Science, University of Auckland
[4]Centre for Brain Research, University of Auckland
[5]Department of Computer Science, Cognitive Science Program, University of Toronto

Abstract

Schemas are mental representations of common structures of our experience, and they are centrally important to human thinking and memory. Recently, it has been proposed that schemas also play an important role in structuring our imagination of the future. However, tools for automatically measuring the schematic content of written and spoken event narratives are underdeveloped. Here, we report a preliminary investigation into a set of metrics that may differentiate between more and less schematic narratives. Across two experiments, we find that written and spoken narratives that are schema-congruent are more *associative*, in that they contain words that are more strongly psychologically associated with one another. We discuss how this finding might contribute to the development of tools to automatically measure schematicity in future narratives.

Keywords: future thinking; narrative; schematicity; schema; word association

Introduction

Humans are perhaps unique in our ability to cast our minds into the future (Suddendorf & Corballis, 2007). This ability, called "episodic future thinking" or EFT (Atance & O'Neill, 2001), is closely related to our ability to remember the past; the "constructive episodic simulation hypothesis" proposes that we imagine the future by recombining perceptual details from memory into novel events by means of so-called "constructive processes" (Schacter & Addis, 2007). Crucially, episodic memory is also thought to rely on a process of recombination and construction rather than strict reproduction (Schacter, 2012). Thus, it has been argued that memory and EFT, rather than being opposites, both reflect the operation of a single "simulation system", the function of which is to mentally simulate events, whether past, future, or atemporal (Addis, 2020). This "continuist" view is contrasted with "discontinuism", which sees memory as a phenomenon fundamentally distinct from imagination and other cognitive processes (Perrin & Michaelian, 2017; Robins, 2020).

Evidence of overlap between memory and EFT comes from neuroimaging studies indicating that the two share the same neural substrates (reviewed in Buckner & Carroll, 2007), as well as from behavioural studies showing that episodic memory impairments, whether due to depression (Addis et al., 2016), post-traumatic stress (Brown et al., 2014), or medial temporal lobe damage (Hassabis et al., 2007), are associated with commensurate impairments in EFT. These latter behaviour studies have relied on textual analysis methods such as the autobiographical interview scoring method (Levine et al., 2002) by which researchers quantify the degree of episodic detail in narratives. Such hand-scoring are somewhat laborious, which has prompted the recent development of automated methods (van Genugten & Schacter, 2022).

In addition to constructive processes, more attention has recently been given to the role of schemas in episodic future thinking (Addis, 2020; Williams et al., 2022). A schema is a high-level knowledge structure that summarizes invariant features of experience (Rumelhart, 1980). The details of this definition, including the degree of theoretical overlap with other constructs such as "gists" and concepts, sometimes differ between research groups (Gilboa & Marlatte, 2017). Nonetheless, significance of schemas has been recognized in memory research since the time of Bartlett (1932), and recent work has explored the important roles schemas play in retrieval, encoding, and even perception (McKenzie et al., 2014; Spalding et al., 2015; Sweegers et al., 2015). Addis (2020, 2018) has proposed that schemas play the same scaffolding role in memory and imagination, guiding mental simulation by providing information about the types of elements likely to be or to have been present and the relationships between them. One prediction of this view is that, when the elements of a simulation are not strongly associated based on past experience, there is less schematic knowledge to draw on to guide event construction, perhaps leading to more subjective mental effort and less detailed mental simulations. However, specific claims like this about the role of schemas in EFT have not been systematically evaluated.

Just as text analysis has been an important tool in studying

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*.

the role of constructive processes in both memory and future thinking (Levine et al., 2002; Renoult et al., 2020), text analysis methods will also be indispensible if we are to expand our understanding of the role of schemas in EFT. Thus, the goal of the current report is to explore of a set of textual features that may be informative about the *schematic content* or *schematicity* of narratives of imagined future events. Identifying such features is a first step toward developing methods for automatically quantifying the schematicity of narratives. Moreover, any tool developed to measure schematicity of future narratives will likely be applicable to memory narratives. These methods would enable research analogous to studies of constructive processes in memory and EFT—i.e., studies identifying the sorts of conditions, lesions, and experimental manipulations that impact the deployment of schemas in imagination. Such studies would, in turn, provide a more complete picture of the neural substrates of schematic knowledge and their role in mental simulations.

Quantifying schematicity of narratives

How might schematicity manifest in text? Previous work in a similar vein has elicited narratives in response to cue words, and has measured schematicity according to the number of words in a narrative related to its cue word (Wynn et al., 2022). Specifically, for each cue, a dictionary of the n most closely-related words was constructed based on similarities according to GloVe embeddings (Pennington et al., 2014), and the overlap between narrative and the dictionary of cue-related words was used as a measure of schematicity. For example, a narrative written in response to the cue word *beach* will contain more words closely related to the word *beach*. This method is something of a generalization of approaches that quantify anxiety, anger, use of stereotypes, or other content in text using dictionaries of words related to the target concept or emotion in question (Tausczik & Pennebaker, 2010; Nicolas et al., 2021).

One drawback of such an approach is that it requires a particular experimental paradigm in which a cue word is presented to participants who then narrate an event related to it. However, cues can be entire phrases (D'Argembeau et al., 2010) or abstract prompts such as time periods (e.g., *3 years from now*) that do not constrain the topic of the narrative Lin & Epstein (2014). Other studies do not use cue words at all (Anderson & Dewhurst, 2009). Thus it would be ideal to have a method of quantifying schemas that are topic-agnostic and do not depend on any particular cueing paradigm.

Moreover, whereas our schemas are based on patterns of co-occurrence in real life, word embeddings such as those provided by the GloVe method are based on patterns of co-occurrence in large text corpora. While the two are certainly not independent, it has been argued that word *associations* (i.e., the degree to which one word calls to mind another) capture the structure of the mental lexicon better than word embeddings based on text corpora (Vankrunkelsven et al., 2021). Indeed, word association data provide better predictions of ratings of the relatedness of words (De Deyne et al., 2015). This may be because obvious information is likely to be left out of most text (e.g., that bananas are yellow), despite the fact that this information is sometimes the basis for a relationship between two words (and for an association between them; Vankrunkelsven et al., 2021). Resources such as ConceptNet (Speer et al., 2017) similarly explicitly spell out relationships between words that might otherwise be taken for granted in normal communication, and in future work we hope to explore the impact of using these resources rather than psychological associations when measuring schematicity. Given some measurement of psychological association between words, how might we use this to estimate the schematicity of a narrative?

One possibility is that a highly schematic narrative includes many words that are strongly associated with one another. This is because both schemas and word associations develop as a result of patterns of co-occurrence in our experience. For example, birthday parties (at least in a North American cultural context) usually involve cake and candles. Thus, our "birthday party" schema includes the associated elements *cake* and *candles*, and the words "cake" and "candles" are associated with one another. A narrative whose construction is strongly guided by a schema will contain many elements of that schema. These will be elements that typically co-occur within a given context, such that the words for these elements will be associated with one another. Thus, we would expect a more schematic narrative to be more *associative*—to contain more words that are strongly associated with one another than a less schematic narrative.

With the advent of large-scale web-based data collection in psychology, it has become possible to obtain normative estimates of associations between large sets of words. The "Small World of Words" association norms (De Deyne et al., 2019) constitute the largest dataset collected to date for this purpose. As its name suggests, the data are suited to a network-based approach, quantifying the association strength between words according to the distance between them in a network whose nodes are words and whose edges are weighted by response frequencies.

To test whether word association metrics might be related to schematicity, we need experimental manipulations that can produce more and less schematic narratives. To that end, we report results from two experiments. Both involve an *incongruent* experimental condition, in which participants imagine and narrate events involving elements that would be unlikely to co-occur in real life. In this condition, participants should not have pre-existing schematic knowledge that would allow for the co-occurrence of these elements and should therefore produce less schematic narratives. Both experiments also employ a *congruent* condition where participants imagine more schema-congruent events and should therefore produce more schematic narratives. In both experiments, participants' self-report ratings of factors such as the imagined events' plausibility were taken as manipulation checks. Nonetheless, these subjective reports cannot be taken as direct measures of schematicity of the type we are seeking. As we will see, despite significant methodological differences between the experiments, a similar pattern of results emerges across both.

Thus, our overall approach (Fig. 1) is to use experimental manipulations (with manipulation checks, as described below) to obtain more schematic (data from the congruent conditions) and less schematic (data from the incongruent conditions) sets of narratives. If our metrics of associativity are higher in the congruent set, they can then be taken as possible indicators of schematicity.

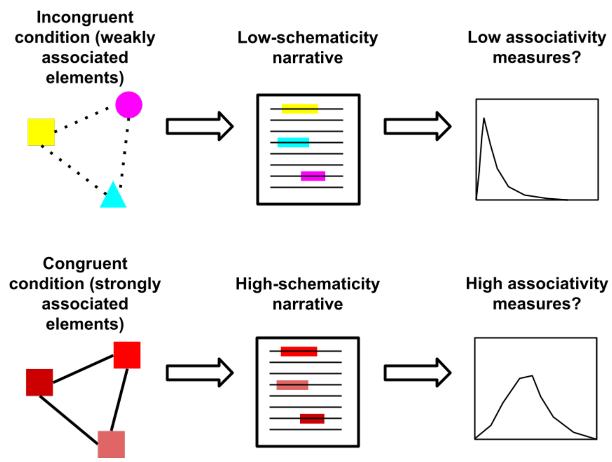

Figure 1: Illustration of the experimental approach. In our experiments, participants constructed narratives in either an incongruent condition, which required imagining a future event containing weakly associated elements (e.g., court room, squeegee, and cookie), or an a congruent condition, which required imagining a future event involving more strongly associated elements (e.g., family home, toy chest, and bathrobe). In the congruent condition, participants have schematic knowledge to support the integration of the elements and should therefore produce more schematic narratives. In the incongruent condition, participants lack this knowledge and should therefore produce less schematic narratives. We can then examine whether these narratives yield different measures of associativity, thereby testing whether associativity is a plausible candidate measure of schematicity for future narratives.

Materials and methods

In this section, we describe our behavioural experiments (one of which was previously published and one of which is novel) and our method of computing associativity as a candidate indicator of schematicity.

Experiment 1

Experiment 1 is described in detail in Roberts et al. (2017) and the present work uses transcript data from that study. Briefly, 33 participants first identified 3 social spheres in their lives (e.g., university) and listed 40 people, objects, and locations encountered within these spheres in the last 5 years with the caveat that each person, object, and location was unique to its given sphere. In a second session, approximately one week later, participants were asked to imagine future events involving a person, object, and location drawn either all from the same sphere (congruent condition; e.g., lecture hall, professor, notebook) or each from a different sphere (incongruent condition; e.g., lecture hall, uncle, soccer ball). They were instructed to verbally describe the events in as much detail as possible during each 3-minute trial, and afterwards they provided ratings of the difficulty of imagining each event, its plausibility, and its similarity to past experiences. Participants also provided the temporal distance of the events (e.g., in 1 week, in 2 months). Each participant provided 4 narratives in each condition and recordings of these verbal narrations were then transcribed. Data from 3 participants was lost due to recorder malfunction and transcripts from an additional 2 narratives could not be located. Thus, in total, 238 narratives were collected from 30 participants.

Experiment 2

60 participants were recruited via Prolific to complete an online study in which, on a series of trials, they were presented with the name of a location and two objects and were asked to write about a possible future event in their lives that would combine all three. On “congruent” trials, the objects presented were ones that would naturally co-occur in the given location, whereas the objects and locations in the “incongruent” condition would be less likely to co-occur. For example, “family home”, “toy chest”, and “bathrobe” were a set of stimuli in the congruent condition and “court room”, “squeegee”, and “cookie” were a set in the incongruent condition. Stimuli in the congruent and incongruent condition were chosen such that they did not significantly differ between conditions in semantic similarity or co-occurrence in text corpora. Each participant completed 3 trials in each condition. These stimuli were validated in a separate study in which participants rated the ease with which they were able to imagine events involving sets of objects and locations. These ratings were then used to construct the congruent and incongruent stimuli in the present study such that events in the congruent trials were easier to imagine than those in incongruent trials. In the present study, participants rated the difficulty of imagining the future events they wrote about, as well as their plausibility and similarity to participants’ past experiences. Participants also provided the temporal distance of the events (e.g., in 1 week, in 2 months). This served in part as a reminder of the instructions to imagine events in the personal future. To ensure that participants did not use generative AI tools to write their responses, their keypress counts were compared to character counts for each narrative. One participant’s data was removed due to a discrepancy in these counts. Data were also removed in 4 cases in which participants skipped a trial without writing any text. Thus, the dataset in experiment 2 comprised 350 narratives from 59 participants.

Incongruent condition

Location: Court room
Object 1: Squeegee
Object 2: Cookie

Narrative: *I am sitting in the courtroom waiting to be judged, with the irritating squeeking of a squagee cleaning the window beside me. This annoys me even more when I can see the cleaner enjoying a cookie while using the squeeger. I am unsure why this would happen in this formal setting of a court room, but I am frustrated and confused.*

Efficiencies (ε):

```
sit - courtroom: 0.0068
courtroom - wait:0.0056
wait - judge:    0.0056
clean - window:  0.02
window - annoy:  0.005
annoy - see:     0.0053
see - cleaner:   0.005
cleaner - enjoy: 0.004
enjoy - cookie:  0.0051
cookie - use:    0.0052
unsure - happen: 0.007
                 ...
```

Congruent condition

Location: Family home
Object 1: Toy chest
Object 2: Bathrobe

Narrative: *I would come back from work into the family home and see the toy chest with some toys left out, which I would gather into it. Then I'd go for a hot shower and enjoy the steam and the heat, after which I'd dry myself and slip on a bathrobe.*

Efficiencies (ε)

```
come - work:     0.0067
work - family:   0.01
family - home:   0.07
home - see:      0.0068
see - toy:       0.0099
toy - chest:     0.01
chest - toy:     0.01
toy - leave:     0.0051
leave - gather:  0.0067
gather - go:     0.0079
go - hot:        0.0173
hot - shower:    0.04
shower - enjoy:  0.005
enjoy - steam:   0.0067
                 ...
```

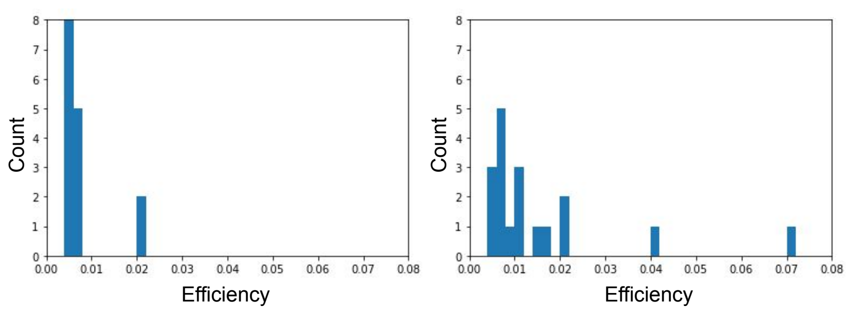

Figure 2: Illustration of the quantification and analytic approach using data from experiment 2. The left panel shows a narrative from the incongruent condition, in which the participant is asked to narrate an event involving 2 objects and a location that would be unlikely to co-occur in real life. As a result, the associativity between tokens is low. In contrast, the narrative in the congruent condition combines elements that could be likely to co-occur in real life and the associativity between tokens is higher.

Quantification of schematicity from text

Preprocessing Tokens were extracted from transcripts using the SpaCy Python library. Only tokens identified by SpaCy's part-of-speech tagger as nouns, verbs, or adjectives were retained, and these were lemmatized and converted to lowercase for further analysis.

Word association network The "Small World of Words" database (De Deyne et al., 2019) provides word association norms from a large set of trials in which participants are presented with a prompt word and provide an open-ended response. From this, we can compute the probability of a response given a prompt, $p(\text{response}|\text{prompt})$, as a measure of the associative strength between the prompt and the response. To construct a network of association strengths, we simply treated each word as a node and set the edge weight between nodes as the associative strength between the relevant words. In cases where words appeared as both prompts and responses, we computed edge weights as

$$W_{\text{word1},\text{word2}} = \max\left(p(\text{word1}|\text{word2}), p(\text{word2}|\text{word1})\right)$$

i.e., as the average of the two directional association strengths. Importantly, we also lemmatized the cues and responses in the Small World of Words dataset and retained only the strongest associations. Thus, for example, the association between *candle* and *birthday* was overwritten by the stronger association between *candles* (lemmatized to *candle*) and *birthday*.

Given such a network, we can define a cost measure as the inverse of the edge weight (i.e., the inverse of associative strength) and obtain the shortest (i.e., minimum cost) path between two words. For example, the shortest path (length 25.85) between the two related words “candles” and “cake” is “candles”, “birthday”, “cake”. In contrast, the shortest path (length 41.58) between the two unrelated words “candles” and “seesaw” is ”candles”, ”light”, ”projector”, ”slide”, ”playground”, ”seesaw”.

For each pair of tokens i and j occurring adjacently[1] in a narrative, we computed the shorted path length $\delta_{i,j}$ and computed the local efficiency of the connection as $\varepsilon_{i,j} = \frac{1}{\delta_{i,j}}$. In the case where there is no path between two words, the local efficiency is 0. Each narrative yielded a distribution of local efficiencies (Fig 2). For each narrative’s distribution, we computed the mean, median, and maximum value as candidate measures of the *associativity* of that narrative (referred to henceforth as “mean associativity”, “median associativity”, and “maximum associativity”).

Results

As a manipulation check, we first compared participants’ self-reported ratings of the difficulty of imagining the events in each condition, their plausibility, and their similarity to past experience. To do so, we computed mixed effects models, modelling participants’ ratings (measured on a 4-point Likert scale in experiment 1 and a 5-point Likert scale in experiment 2) with a fixed effect of experimental condition (congruent versus incongruent) and a random intercept for each participant. In both experiments, participants rated the events in the incongruent condition as more difficult to imagine than those in the congruent condition ($p < 0.001$ in both cases), as less plausible ($p < 0.001$ in experiment 1; $p = .019$ in experiment 2), and as less similar to their past experiences ($p < 0.001$ in experiment 1; $p = .010$ in experiment 2).

To compare measures of associativity between conditions in each experiment, we again computed mixed effects models with a fixed effect of experimental condition and a random intercept for each participant. This was intended to account for the possibility that different participants might produce narratives that differed systematically in their schematicity. In experiment 1 (Fig. 3, top), we found significantly greater mean and median associativity in the congruent compared to the incongruent condition ($p = .031$ and $p = .004$, respectively), and a borderline-significant difference in maximum associativity ($p = .053$). In experiment 2 (Fig. 3, bottom), we found significantly greater mean and maximum associativity in the congruent compared to the incongruent condition ($p = .013$ and $p = .012$, respectively), but no statistical difference in median associativity ($p = .382$).

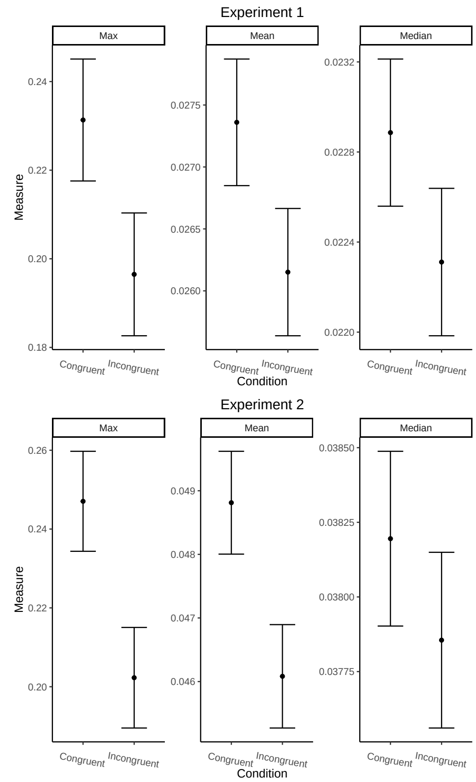

Figure 3: Fixed effects of the congruent versus incongruent experimental conditions for each measure of associativity across experiments 1 (top panel) and 2 (bottom panel). Error bars reflect one standard error of the fixed effect estimate.

We next explored the results on an individual basis: we binarized the mean associativity measure on an individual-wise basis by averaging the measures within each condition for each participant and examining whether the measure was higher in the congruent than in the incongruent condition. Across the two experiments, this was true in 54 out of 89 cases (Fig 4; $p = .028$ according to a binomial test)[2].

To compare the measures of associativity across the two experiments, we computed another mixed effects model with fixed effects of both congruency and experiment, and a random intercept for each participant. The congruency effect was more significant in this pooled analysis than in either of the experiments considered individually ($p = .003$). Interestingly, there was also a highly significant effect of the experiment ($p < 0.001$) such that mean associativity in the

[1] We repeated the same analysis using all possible pairs of tokens in the narrative but found strong correlations with the measures computed using only adjacent tokens. We report the results using adjacent tokens here because the computational savings of this approach are considerable.

[2] However, this difference was not significant when the experiments were considered individually (experiment 1: 19 out of 30 cases, $p = .100$; experiment 2: 35 out of 59 cases, $p = .096$)

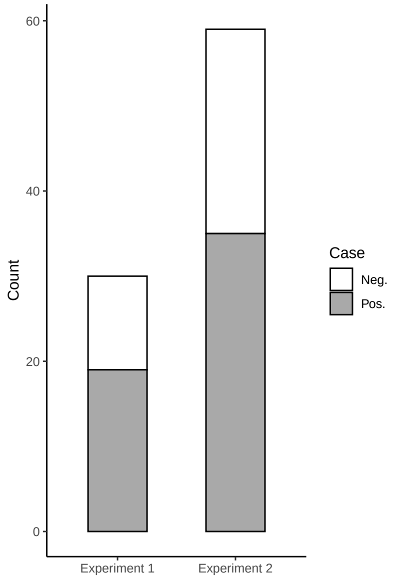

Figure 4: Counts of individual participants whose average measures of schematicity between conditions differed in the expected versus unexpected directions across both experiments. "Positive" cases are those in which the participant's average schematicity was higher in the congruent versus incongruent condition, and "negative" cases are those in which the opposite pattern held.

second experiment (involving written narratives) was higher than those in the first experiment (involving transcribed spoken narratives).

It seemed possible that this was a byproduct of the different narrative lengths between the two experiments (narratives tended to be far longer in the first experiment than then second—1601 characters on average vs. 445 characters; $t = 25.42$, $p < 0.001$). Thus, we added to the mixed effect model on the pooled data a fixed effect of narrative length (measured in number of characters). However, this effect was not significant ($p = .672$). Thus, the different measures of associativity produced by the two experiments may have been due to differences in methodology (written versus spoken narratives; experimenter- versus participant-provided cues; or autobiographical specificity of schemas) rather than a result of differences in narrative lengths.

Discussion

The present work aimed to explore a set of metrics that may differentiate between more and less schematic narratives with an eye toward developing a measure of schematicity. How might the measures of associativity presented here be improved so that they differentiate more reliably between more and less schematic narratives, and thereby represent more credible candidate measures of schematicity? First, there are technical improvements that could be made; for example, it is not always clear when a compound word should be treated as two words (e.g., *courtroom* versus *court room*; Fig. 2) or vice versa.

Another potential issue with the current method is that it might sometimes overestimate the relatedness of two concepts (though we have not systematically investigated how often this occurs): two words can sometimes be connected through two different senses of a third word. For example, using the method presented here, the shortest path between "page" and "tree" is "page", "leaf", "tree"—presumably because one leafs through the pages, or leaves, of a book, and because trees have leaves. Similarly, the shortest path between "job" and "light" is "job", "fire", "light"—one is fired from a job and fires produce light. Connections such as these are spurious in that they are not based on the real-life co-occurrence of two concepts, actions, objects, etc. and thus may be misleading as to the contents of people's schemas. Thus, the current measures might be refined by exploring alternative methods of quantifying the relatedness of words. One such method is based on the degree of overlap in responses between two cue words (De Deyne et al., 2019). Although this limits the analysis to those words presented as cues, it may provide more reliable measures of relatedness by considering the entire set of responses to cue words rather than seizing on a single (possibly spurious) connection.

Nonetheless, the broad approach presented here seems to hold promise. For example, in experiment 1, the incongruency of the elements in the incongruent condition was based on participants' own particular experiences—that is, the elements were unlikely to co-occur in participants' own lives. However, in principle, this may not manifest in differences in word associations, which are based on the shared experiences of many individuals. To illustrate: within my life, it may be hard to imagine finding myself in my classroom with my red notebook, which I use as a personal journal. Nonetheless, the words "notebook" and "classroom" are in general associated—as it happens, through the shortest path "notebook", "binder", "teacher", "classroom" (length 23.52). In other words, in experiment 1, participants imagined events that were congruent or incongruent with schemas they had developed based on their *particular* experiences, whereas it seems that word associations would be best suited to capturing *shared* schemas based on experiences shared by many individuals. Thus, it is an encouraging sign that the word association-based metrics did in fact differ between the congruent and incongruent conditions in experiment 1. It may be that, in the incongruent condition, participants had to construct highly unlikely scenarios in order to explain the co-occurrence of the incongruent elements, and that these scenarios were incongruent not only with their own schemas but with the shared schemas captured by word associations. Whether it is generally true that we depart from shared schemas when we depart from our idiosyncratic schemas remains to be investigated in future research. If so, the outlook is positive for an automated, general-purpose measure of narrative schematicity based on word associations.

Acknowledgments

This work was supported by the Data Sciences Institute at the University of Toronto, and by a Discovery Grant (RGPIN-2023-04644) from the Natural Sciences and Engineering Research Council of Canada. DRA is supported by the Canada 150 Research Chairs Program.

References

Addis, D. R. (2018). Are episodic memories special? on the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand*, *48*(2-3), 64–88.

Addis, D. R. (2020). Mental time travel? a neurocognitive model of event simulation. *Review of Philosophy and Psychology*, *11*, 233–259.

Addis, D. R., Hach, S., & Tippett, L. J. (2016). Do strategic processes contribute to the specificity of future simulation in depression? *British Journal of Clinical Psychology*, *55*(2), 167–186.

Anderson, R. J., & Dewhurst, S. A. (2009). Remembering the past and imagining the future: Differences in event specificity of spontaneously generated thought. *Memory*, *17*(4), 367–373.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in cognitive sciences*, *5*(12), 533–539.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge university press.

Brown, A. D., Addis, D. R., Romano, T. A., Marmar, C. R., Bryant, R. A., Hirst, W., & Schacter, D. L. (2014). Episodic and semantic components of autobiographical memories and imagined future events in post-traumatic stress disorder. *Memory*, *22*(6), 595–604.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in cognitive sciences*, *11*(2), 49–57.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "small world of words" english word association norms for over 12,000 cue words. *Behavior research methods*, *51*, 987–1006.

De Deyne, S., Verheyen, S., & Storms, G. (2015). The role of corpus size and syntax in deriving lexico-semantic representations for a wide range of concepts. *Quarterly Journal of Experimental Psychology*, *68*(8), 1643–1664.

D'Argembeau, A., Ortoleva, C., Jumentier, S., & Van der Linden, M. (2010). Component processes underlying future thinking. *Memory & cognition*, *38*(6), 809–819.

Gilboa, A., & Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends in cognitive sciences*, *21*(8), 618–631.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, *104*(5), 1726–1731.

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychology and aging*, *17*(4), 677.

Lin, H., & Epstein, L. H. (2014). Living in the moment: effects of time perspective and emotional valence of episodic thinking on delay discounting. *Behavioral neuroscience*, *128*(1), 12.

McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., Rivière, P. D., & Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron*, *83*(1), 202–215.

Nicolas, G., Bai, X., & Fiske, S. T. (2021). Comprehensive stereotype content dictionaries using a semi-automated method. *European Journal of Social Psychology*, *51*(1), 178–196.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (emnlp)* (pp. 1532–1543).

Perrin, D., & Michaelian, K. (2017). Memory as mental time travel. *The Routledge handbook of philosophy of memory*, 228–239.

Renoult, L., Armson, M. J., Diamond, N. B., Fan, C. L., Jeyakumar, N., Levesque, L., ... others (2020). Classification of general and personal semantic details in the autobiographical interview. *Neuropsychologia*, *144*, 107501.

Roberts, R., Wiebels, K., Sumner, R., van Mulukom, V., Grady, C., Schacter, D., & Addis, D. (2017). An fmri investigation of the relationship between future imagination and cognitive flexibility. *Neuropsychologia*, *95*, 156–172.

Robins, S. (2020). Defending discontinuism, naturally. *Review of Philosophy and Psychology*, *11*(2), 469–486.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In *Theoretical issues in reading comprehension* (pp. 33–58). Routledge.

Schacter, D. L. (2012). Constructive memory: past and future. *Dialogues in clinical neuroscience*, *14*(1), 7–18.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 773–786.

Spalding, K. N., Jones, S. H., Duff, M. C., Tranel, D., & Warren, D. E. (2015). Investigating the neural correlates of schemas: ventromedial prefrontal cortex is necessary for normal schematic influence on memory. *Journal of Neuroscience*, *35*(47), 15746–15751.

Speer, R., Chin, J., & Havasi, C. (2017). Conceptnet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 31).

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and brain sciences*, *30*(3), 299–313.

Sweegers, C., Coleman, G., Van Poppel, E., Cox, R., & Talamini, L. (2015). Mental schemas hamper memory storage of goal-irrelevant information. *Frontiers in human neuroscience*, *9*, 629.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24–54.

van Genugten, R., & Schacter, D. L. (2022). Automated scoring of the autobiographical interview with natural language processing. *PsyArXiv*.

Vankrunkelsven, H., Vankelecom, L., Storms, G., De Deyne, S., Voorspoels, W., Kristiansen, G., ... Zhang, W. (2021). Guessing words: A comparison of text corpus and word association models as the basis for the mental lexicon. In *Cognitive sociolinguistics revisited.* De Gruyter Mouton.

Williams, K., Snytte, J., & Sheldon, S. (2022). Individual differences in depression are reflected in negative self-evaluations when imagining future events. *Journal of Behavioral and Cognitive Therapy*, *32*(3), 207–221.

Wynn, J. S., Van Genugten, R. D., Sheldon, S., & Schacter, D. L. (2022). Schema-related eye movements support episodic simulation. *Consciousness and cognition*, *100*, 103302.