











Temporal Contiguity in Incidentally Encoded Memories

M. Karl Healey
Michigan State University

The law of similarity is a key principle of memory: recalling one memory tends to trigger recall of similar memories. The amount of time separating two events is believed to be a powerful determinant of similarity. If so, there should be a strong Temporal Contiguity Effect whereby thinking of one event triggers  of other events experienced nearby in time.  Evidence for this Temporal Contiguity Effect has come almost exclusively from tasks that explicitly ask subjects to memorize items. Are memories naturally organized by temporal proximity when there is no intent to memorize, as is often the case outside the  Two experiments show that removing the intent to memorize can also  moves the Temporal Contiguity Effect.  Third experiment shows that it is not the intent  to memorize per se, but rather how subjects process information while incidentally learning it that generates temporal contiguity.



Keywords: episodic memory; free recall; temporal contiguity

Success in remembering important information is powerfully influenced by how we encode that information (Craik & Lockhart, 1972; Tulving & Pearlstone, 1966). Control processes (Lehman & Malmberg, 2013; Raaijmakers & Shiffrin, 1981) allow us to strategically choose how we process information so  to maximize recall (e.g., Unsworth, 2016).  We know surprisingly little about how control processes impact one of the oldest and most influential observations about memory: Recalling one event tends to trigger recall of other events experienced nearby in time (Aristotle, n.d.; Bower, 1972; Kahana, 1996).

This Temporal Contiguity Effect (TCE) manifests in many laboratory tasks (Davis, Geller, Rizzuto, & Kahana, 2008; Kahana, 1996; Schwartz, Howard, Jing, & Kahana, 2005). For example, in free recall you study a list of words presented serially and recall them in any order. If you have just recalled the word studied in the 5th serial position, the  word you recall is much more likely to be from the 6th  position than a more distant position (Kahana, 1996).

The TCE has shaped theories of episodic memory (Cortis, Dent, Kennett, & Ward, 2015; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Lehman & Malmberg, 2013; Lohnas, Polyn, & Kahana, 2015; Unsworth, 2008), the testing effect (Karpicke, Lehman, &

Aue, 2014), directed forgetting (Sahakyan, Delaney, Foster, & Abushanab, 2013), retrieval induced forgetting (Kliegl & Bäuml, 2016), childhood development (Jarrold et al., 2015), cognitive aging (Healey & Kahana, 2016; Wahlheim & Huff, 2015), event segmentation (Ezzyat & Davachi, 2014), time estimation (Sahakyan & Smith, 2014), and even visual perception (Turk-Browne, Simon, & Sederberg, 2012). Yet, little is known about how encoding control processes influence contiguity (Hintzman, 2016).

Because control processes are deliberately engaged to aid encoding, their effect on the TCE can be studied by comparing explicit encoding conditions in which subjects  to memorize items with  incidental encoding conditions in which subjects do not suspect their memory will be tested and thus have no impetus to engage encoding control processes. Theories of the TCE make very different predictions about what should happen under incidental encoding conditions. Some models suggest the TCE does not depend directly on control processes and should be easily observed under incidental encoding (Healey, Crutchley, & Kahana, 2014; Lohnas et al., 2015). Others suggest that the TCE is entirely due to task-specific strategies implemented by control processes and should disappear under incidental encoding (Hintzman, 2016).

These predictions have never been directly tested. The only study to measure the TCE after incidental encoding (Nairne, Cogdill, & Lehman, submitted) did not include an explicit encoding control condition because it was focused on the survival processing effect (Nairne & Pandeirada, in press), not the role of encoding intent. Tantalizingly, this study did *not* find a significant TCE, but the lack of a control condition leaves it unclear whether it was the absence of intent or some other factor that eliminated the TCE. Therefore,

M. K. Healey is the sole author of this article and is responsible for its content.

phone: 517-432-3107

Version of November 2, 2016

1050 Words excluding Method and Results (Approximate due to use of L^AT_EX)

Table 1
Sample sizes, exclusions, and recall probability by condition.

Exp	Condition	n Included	n A	Recall Prob. (SD)
Exp1	Explicit	290	–	0.46 (0.16)
	Incidental	339	48	0.40 (0.16)
Exp2	Explicit	125	–	0.47 (0.17)
	Incidental	212	44	0.47 (0.14)
Exp3	Weight	291	34	0.42 (0.15)
	Animacy	294	33	0.37 (0.14)
	Scenario	299	33	0.38 (0.14)
	Movie	329	59	0.38 (0.15)
	Relational	311	102	0.44 (0.18)

Experiment 1 directly compares the TCE under explicit and incidental encoding.

Experiment 1

Method

Subjects

Given that manipulating encoding intention might reduce, but not eliminate, the TCE, it is critical to have sufficient power to detect a small effect. Sederberg, Miller, Howard, and Kahana (2010) reported a meta-analysis of the TCE in explicit encoding studies; power calculations reveal that a sample size of 143 per condition would provide a $1 - \beta$ power of 0.95 to detect (via a 1-tailed 1-sample t-test) an effect one fifth the size of the average effect they reported. To collect enough data to meet or exceed this sample size in the Incidental conditions, subjects for all of the experiments reported here were recruited using Amazon's Mechanical Turk, a crowdsourcing website that allows for efficient collection of large volumes of high-quality data. Subjects were paid \$10 for participating.

Subjects in the incidental memory condition were excluded from analysis if they reported in a post-experiment questionnaire that they suspected that their memory would be tested while they were performing the judgment task. The final analyzed sample was composed of 290 in the explicit condition and 339 in the incidental condition. Table 1 shows the total number of included and excluded subjects for each experiment.

Procedure

All subjects completed two free recall lists (only the first of which is analyzed here). Each list was composed of 16 words drawn randomly from a pool of 1638 words, with

the constraint that no word was used more than once for a given subject. Words were presented one at a time on the subject's computer screen for 4 seconds. There was an inter-stimulus interval of 1 second between word presentations during which a fixation cross was displayed in the same location in which the words appeared. The final word of each list was followed by a 16 seconds distractor period during which subjects answered math problems of the form $A + B + C = ?$ where A , B , and C were positive, single-digit integers, though the answer could have been one or two digits. Subjects typed their answers to the math problems in a text-box and pressed enter to submit the answer. Upon pressing enter, a new math problem and a new blank text-box appeared. Subjects were instructed to "Try to solve as many problems as you can without sacrificing accuracy. The task will automatically advance when the time is up."

Following the math distractor task, subjects in both conditions were asked to recall as many items as possible from the preceding list, in any order. Subjects typed each recalled word in a text-box and pressed enter to submit the word. Upon pressing enter, the word disappeared and a new blank text-box appeared such that subjects could not see their prior responses. Subjects were given 75 seconds to recall as many words as they could. To ensure subjects noticed that the recall period had begun (e.g., were not looking at the keyboard and typing their answer from the final math problem), a red screen was displayed for 500 ms before the recall instructions were displayed and the recall text-box did not begin accepting input for a further 500 ms. Therefore, including the math distractor, there was a total of $16 + 0.5 + 0.5 = 17$ seconds delay between the end of the study period and the beginning of the recall period.

Encoding Instructions Manipulation. Subjects were randomly assigned to either the Incidental condition or the Explicit condition. Prior to seeing the first list, subjects in both conditions were told that they would see a series of words and make a simple judgment about each one (whether it fit in a shoebox). The exact instructions depended on the condition. In the Explicit condition, subjects were given standard free recall instructions that described the size judgment task but emphasized memory. Because the wording of the instructions are integral to the intent manipulation, they are quoted directly here:

Explicit Instructions:

Thank you for participating in this study.

We are interested in how people make simple judgments about common words and how they subsequently remember the words. Please position this window in the center of your screen so you can comfortably view the words.

You will see a list of words appear one at a time and make a judgment about each one (more de-

tails on the next page). After the list of words, you will do a few math problems. After the math problems, you will be prompted to type in any words that you can remember from the list.

When prompted, type any words that you can remember from the list you just saw, *in any order* (type one word in each of the provided text boxes).

[a screen showing task instructions, which did not mention memory and was identical for both conditions]

Your main task is to remember as many of the words as possible; at the end of the list you will be prompted to type as many words as you can remember from the list you just saw, *in any order*.

Implicit Instructions:

Thank you for participating in this study.

We are interested in how people make simple judgments about common words. Please position this window in the center of your screen so you can comfortably view the words.

You will see a list of words appear on the screen and make a judgment about each one (more details on the next page). After the list of words, you will do a few math problems.

[a screen showing task instructions, which did not mention memory and was identical for both conditions]

Your main task is to make as accurate a judgment as possible about each word.

Shoebox Task. In both conditions subjects were asked to make a size judgment about each word while it was present on the screen. Specifically, they were asked to judge if the word referred to an object that would fit into a shoebox. To allow for the same yes/no response across all the encoding tasks used in in all Experiments, subjects were told that we were norming the words for a later study to find items that were neither too easy nor too hard to process and they should press “Y” if it was easy to judge if the item would fit in a shoebox and “N” if it was difficult to make the judgment. See the supplemental materials for the exact task instructions.

Because subjects completed the task online and could not ask an experimenter for clarification, several measures were taken to ensure that subjects understood how to make a response and could be confident that their responses were being registered: During presentation of the lists, a task prompt was displayed above each word (e.g., “Is it easy to judge if it would fit in a shoebox?”). An instruction about how

to make a response was displayed below each word (“Y” for yes, “N” for no”). The task prompt and response instructions were in lighter gray font than the black font used for words. The prompts disappeared once the subject made a valid response, but if the subject made an invalid response (e.g., pressing “B” instead of “Y” or “N”) the response instructions were replaced with an error message in red font until a valid response was made. The word remained on screen for the full 4 second presentation period regardless of the subject’s response.

Recall Scoring

Because subjects typed their responses, typos are likely and counting only exact matches with list words as correct would underestimate their recall scores. Therefore, a typo-sensitive scoring algorithm was implemented as follows: Subject’s responses were converted to lower case and stripped of any white space. Next, each response was compared to all the list-words that the subject has seen so far in the experiment (which were also lowercase and free of white space). If the response exactly matched any of these presented words, it is scored as a correct recall or a prior list intrusion depending on whether the matching list-word was presented on the current or a previous list. If the response did not exactly match a presented word, it was compared with each of the 235886 words in Webster’s Second International dictionary (<https://libraries.io/npm/web2a>). If the response exactly matched any word in the dictionary, it is scored as an extra-list intrusion. If the response did not exactly match any word in the dictionary, it was assumed to be a typo and an attempt was made to correct its spelling.

The spell-checking algorithm began by computing the Damerau-Levenshtein distance Damerau (1964) between the response and each word in the dictionary, providing a measure of the response’s similarity to each candidate word. Because almost all responses in free recall correspond to words that were presented on some list (i.e., extra-list intrusions are rare), the algorithm did not automatically replace the mistyped response with the most similar word in the dictionary. Instead it found the shortest distance between the response and an actually presented list-word, and then found where this “nearest list-neighbor” distance lay in the distribution of distances between the response and the dictionary words. If the nearest list-neighbor distance was below the tenth percentile of the distribution (i.e., if the response was closer to a list item than it was to 90% of the words in the dictionary) it was assumed to be that list item, otherwise it is assumed to be an ELI.

Quantifying the Temporal Contiguity Effect

The TCE is most often examined using a *lag conditional-response probability* function or lag-CRP. The lag-CRP gives the probability that recall of an item studied in position *i* of

a study list will be followed by recall of an item studied in position $i + \text{lag}$. For example, if recall of the item from position 5 was followed by recall of the item from position 6 the lag would be 1. If recall of the item from position 3, the lag would be -2. For each lag, the CRP is computed by dividing the number of times a transition of that lag was *actually* made by the number of times it *could* have been made (e.g., it could not have been made if the item $i + \text{lag}$ was already recalled; Kahana, 1996). The lag-CRP is typically highest for lags 1 and -1 and decreases sharply for larger absolute values of lag. If the TCE is reduced under incidental encoding conditions, the lag-CRP should be flatter.

The lag-CRP provides a visual representation of the TCE, but it is useful to have a single number that quantifies the size of the effect. For this purpose, the *temporal factor score* is typically used (Polyn, Norman, & Kahana, 2009; Sederberg et al., 2010). The temporal factor score is computed by ranking the absolute value of the lag of each actual transition with respect to the absolute values of the lags of all transitions that were possible at that time, which provides a percentile score for each transition. Averaging these percentile scores across all of a subject's transitions provides the temporal factor score.

When evaluating the size of the temporal contiguity effect, it is important to take into account the fact that some items are more likely to be recalled than others and that the likelihood of successful recall is not random with respect to serial position (e.g., primacy and recency effects, or more generally autocorrelations in goodness of encoding). As a consequence, temporally adjacent items will tend to have similarly high or similarly low probabilities of being recalled, creating temporally isolated "pockets" of recallable items, which can influence the size of the TCE (Healey, Long, & Kahana, 2010; Hintzman, 2016; Sederberg et al., 2010). For example, if items from the beginning of the list are more likely to be successfully encoded than items from the middle of the list (i.e., a primacy effect), short-lag transitions between primacy items will naturally be more likely than long-lag transitions between primacy and mid-list items (because you can only transition between items that have been encoded), artificially increasing the TCE. In other words, even if recall involved randomly selecting from the pool of successfully encoded items, completely ignoring lag, one would expect a small TCE.

The size of this artificial TCE can be estimated by taking the items which a subject actually recalled for a given list, randomly shuffling (i.e., permuting) the order of recalls, and recomputing the temporal factor score. Repeating this permutation procedure many times provides a distribution of the temporal factor score expected if recall transitions are completely random with respect to lag. This logic was used to provide a corrected measure of the TCE for each participant.

For each list, the temporal factor score was computed for the actual recall sequence and for 10,000 random permutations of the sequence. The actual temporal factor score was then converted into a z-score, $Z(\text{TCE})$, by subtracting the permutation distribution's mean and dividing by its standard deviation. In the absence of a true TCE, the expected value of $Z(\text{TCE})$ is zero, so we can test for a TCE by determining if the across-subject average of $Z(\text{TCE})$ is significantly above zero.

Data Sharing

All data analyzed in this report will be made freely available on the author's website (cbcc.psy.msu.edu/data)

Results

Figure 1 shows the lag-CRP and corrected temporal factor scores for the Explicit and Incidental conditions. The Explicit condition shows a clear TCE: the lag-CRP is highest for short lags ($|\text{lag}| = 1$) and decreases for larger lags. Moreover, the 95% confidence interval on the $Z(\text{TCE})$ lies well above zero. By contrast, the Incidental condition shows no evidence of a TCE: the lag-CRP is nearly flat and the 95% confidence interval on the $Z(\text{TCE})$ includes zero. Overall recall probability (Table 1) was also lower in the Incidental condition.

These results suggest that removing the intent to encode eliminates the temporal contiguity effect. Because the TCE has proven to be so robust, one can be justifiably skeptical of a single experiment showing lack of contiguity. Therefore, we attempt to replicate the finding in Experiment 2 using a slightly different processing task.

Experiment 2

Method

The methods were identical to those used in Experiment 1 except the judgment task instructions (see Table 1 for sample size information).

The processing required by the *bebox* Task from Experiment 1 is quite simple. So simple that one could argue it is ineffective at forming strong memories, which may artificially reduce the temporal contiguity effect. Therefore, we wanted to retain the basic task of judging size while increasing memory performance in the Incidental condition. That is, can processing that promotes memory do so without producing substantial contiguity? Two effective ways to improve memory mental imagery and self-referential processing. Thus, the Front Door Judgment task asked subjects to imagine trying to move the object referred to by each item through the front door of their house, and decide whether or not it would be possible (again, subjects were asked to indicate if this judgment was easy or difficult to make by pressing

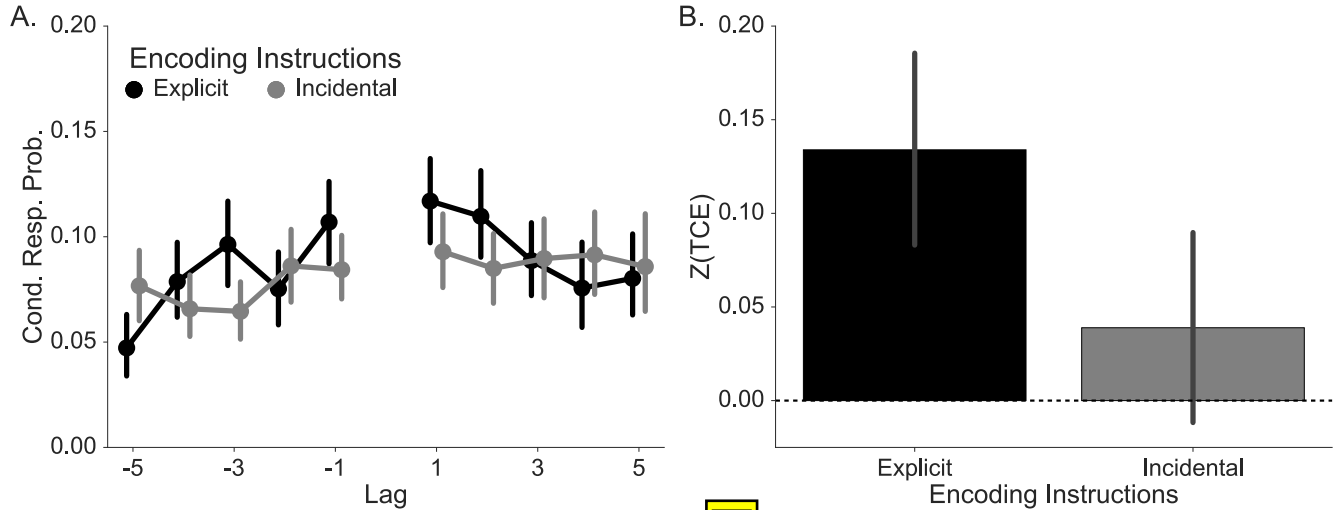


Figure 1. The temporal contiguity effect (TCE) with the Shoebox judgment task under explicit versus incidental encoding. (A) Lag-conditional response probability functions. Error bars are bootstrapped within-subject 95% confidence intervals. (B) The average $Z(\text{TCE})$. Error bars are bootstrapped between-subject 95% confidence intervals. $Z(\text{TCE})$ for a given subject is computed as follows: An observed temporal factor score was computed as the average percentile ranking the temporal lag of each actual transition in the recall sequence with respect to the lags of all transitions that were possible at that time. To determine the temporal factor score expected by chance, a permutation distribution was created by randomly shuffling the order of recalls within the sequence 10,000 times and computing a temporal factor score for each shuffling. The reported value, $Z(\text{TCE})$, is z-score of the observed temporal factor score within the permutation distribution.

or “See the supplemental materials for the exact task instructions.

Results

As predicted, the Front Door task substantially improved memory accuracy in the Incidental condition to approximately the level seen in the Explicit condition of Experiment 1 (Table 1). The Front Door task did not, however, produce a significant TCE. Figure 1 shows that whereas the Explicit condition showed a distinctly peaked lag-CRP and a $Z(\text{TCE})$ significantly above zero, the Incidental condition showed a flattened lag-CRP and a $Z(\text{TCE})$ for which the confidence interval included zero.

Interim Discussion

Experiments 1 and 2 show that the TCE can be absent when intent to encode is absent. One interpretation is that under Incidental encoding, memories are formed by automatic processes similar to those that operate outside the laboratory when we form new memories without trying to memorize anything, and that these processes do not generate contiguity. Explicit encoding adds an additional set of strategic or control processes, which do generate contiguity. Under this interpretation, the contiguity-generating processes are more or less inseparable from the intent to encode. An alternate interpretation is that automatic encoding processes do indeed

tend to produce contiguity, but their effect is obscured by additional, non-memory, processes required by the judgment task. Under this interpretation, the effect of explicit control processes is not to directly produce contiguity but to compensate for the contiguity-obscuring processes in the Incidental condition.

For example, most models produce a TCE because the representations of items studied close together are more similar to each other than they are to the representations of items studied far apart (Healey et al., 2016). But the Shoebox and Door tasks encourage subjects to maintain a common mental representation (an image of a shoebox, or their home) through the list presentation. If this representation is incorporated into the representations of list items, it would increase the similarity of items separated by short lags, attenuating the TCE. In the Explicit condition, subjects likely engage in processing of the items beyond that needed to make a size judgment, which might decrease the similarity of successive items, mitigating the effect of associating each item with a common representation. Although speculative, this is a plausible example of how the judgment task might decrease the TCE in a way that is not due to the lack of intent to encode.

More generally, if automatic encoding processes do not generate contiguity but control processes do, it should be very difficult, or perhaps impossible, to observe a TCE under incidental encoding. But if contiguity is simply obscured by some types of task-related processing, it should be very easy

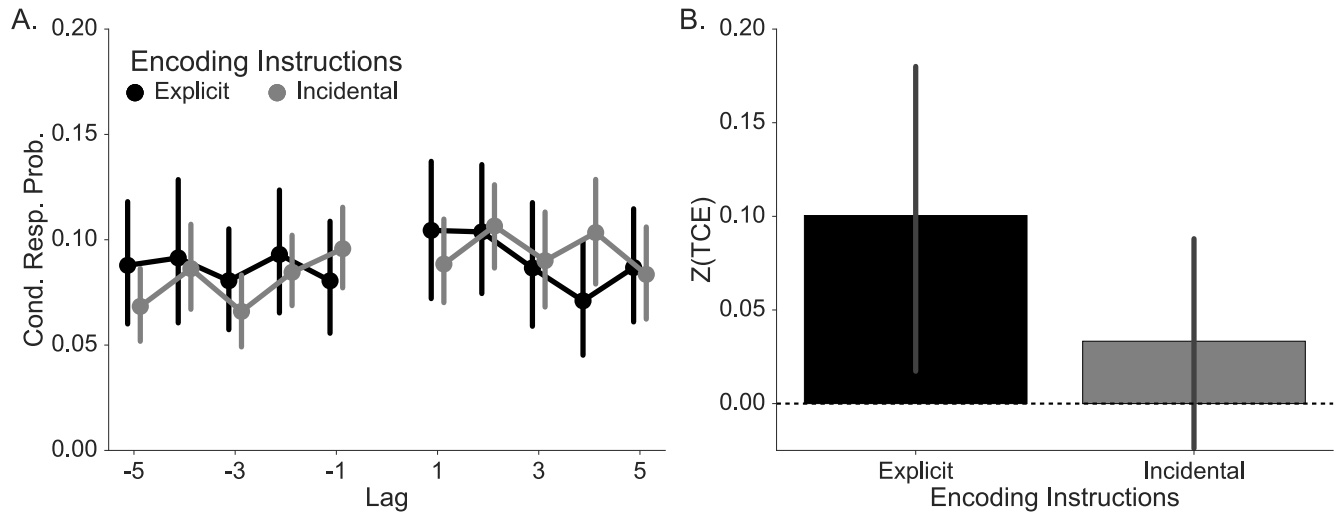


Figure 2. The temporal contiguity effect (TCE) with the Front Door size judgment task under explicit versus incidental encoding. (A) Lag-conditional response probability functions. Error bars are bootstrapped within-subject 95% confidence intervals. (B) The average $Z(\text{TCE})$. Error bars are bootstrapped between-subject 95% confidence intervals. $Z(\text{TCE})$ for a given subject is computed as follows: An observed temporal factor score was computed as the average percentile ranking the temporal lag of each actual transition in the recall sequence with respect to the lags of all transitions that were possible at that time. To determine the temporal factor score expected by chance, a permutation distribution was created by randomly shuffling the order of recalls within the sequence 10,000 times and computing a temporal factor score for each shuffling. The reported value, $Z(\text{TCE})$, is z-score of the observed temporal factor score within the permutation distribution.

to find incidental encoding tasks that produce a TCE. Experiment 3 tests these predictions by examining five different encoding tasks that differ in the type of processing required.

Experiment 3

Method

The question is no longer whether Explicit encoding produces a larger TCE than Incidental encoding, but rather whether the TCE can ever be observed under Incidental encoding. Thus, in Experiment 3 all subjects were given Incidental encoding instructions, but were randomly assigned to one of five different judgment tasks that varied in the type of processing required. Otherwise, the methods were identical to those used in Experiments 1 and 2 (see Table 1 for sample size information).

Processing Task Manipulation

In all conditions, subjects were asked to make a judgment about each word as it was presented. Here we describe the type of processing that each task was intended to discourage (or encourage). Again, to allow for the same yes/no response for each task, subjects were asked to indicate if the judgment was easy to make under the guise of norming the items for a later study. See the supplemental materials for the exact task instructions.

Weight Task. The Weight Task was similar to the size judgment tasks used in the first two experiments except that it asked subjects to compare each item's *weight* to a common referent: a bottle of water. Specifically, they were asked to judge whether each word referred to an object that was heavier than "a standard bottle of water you'd purchase from a vending machine". Because weight is not an easily visualizable attribute, the Weight Task might be expected to reduce the likelihood that subjects will maintain a vivid mental image throughout the list. Thus, it may produce a larger TCE if associating each item with a common mental image tends to attenuate the TCE.

Animacy Task. The Animacy Task asks subjects whether each item refers to an object that is living or non-living. Like the size and Weight tasks, the Animacy task requires subjects to consider only a single attribute of each item (animacy status). But unlike the size tasks, it does not provide a reference object against which to compare each item. Thus, it further reduces the likelihood of maintaining a single vivid image throughout the list.

Scenario Task. The Scenario Task asks subjects to judge the relevance of each word to a scenario; moving to a foreign land (Nairne et al., submitted). Subjects are asked to maintain some representation of this scenario across items, but because it does not specify any pre-existing dimension, like size or weight, each item may be expected to activate many different attributes, lowering the similarity of mental representations from item to item.

Movie Task. The instructions for the Movie Task explain that “when you read a word, it can trigger many different thoughts” and gives the example of the word baseball triggering a series of thoughts: “you might have a mental image of a baseball, you might hear the crack of a bat hitting a ball, you might think of related concepts like ballpark, players, and fans...”. It then asks subjects to allow each item “to activate as many different thoughts as possible. Then use these thoughts to generate a mental movie (like a detailed image of spending an afternoon at a baseball game or what it is like to be a player on a baseball field).” Subjects then judge whether or not it was easy to form such a mental movie. This task removes the requirement to consider each item along the same dimensions and instead encourages subjects to think deeply about the unique attributes of each item, which might be expected to cause very different mental representations to be activated with each successive item, perhaps increasing the TCE.

Relational Task. The Relational task is similar to the Movie Task except instead of being asked to make a new mental movie for each item, subjects are asked to integrate each new item into an ongoing movie: “try to incorporate each new word into your existing mental movie. For example, if the next word was “owner”, you should allow it to activate many associated thoughts and then incorporate it into your existing “ballpark” movie.” By explicitly asking subjects to form links between temporally proximate items, this condition should maximize the chance of observing a TCE.

Results

As seen in Figure 3, all of the processing tasks produced a TCE under incidental encoding conditions. For each task, the lag-CRP tends to decrease with increasing $|lag|$ and the $Z(TCE)$ is significantly above zero. These results show that while the TCE can be attenuated under incidental conditions (as in Experiments 1 and 2), the lack of intent to encode, per se, does not eliminate contiguity.

Indeed, perhaps the most remarkable feature of the data is how little the size of the TCE differs among the tasks, as would be expected if the TCE is due to automatic encoding processes that are independent of the judgment task. The only condition for which the $z(TCE)$ differed significantly from any other condition was the Relational Task condition, which asked subjects to integrate each item in to an ongoing movie. In other words, robust contiguity of approximately the same magnitude was observed regardless of processing task unless subjects were explicitly encouraged to integrate the current item with past items.

General Discussion

Experiment 1 showed that the TCE can be eliminated under incidental encoding. Experiment 2 replicated this finding. In both of these experiments, incidental encoding task

encouraged keeping a specific visual image active through the list, possibly increasing the similarity of memories encoded far apart in time. Experiment 3 used five different implicit encoding tasks that did not share this feature and found a clear TCE in each case. That is, these results are an existence proof for temporal contiguity under incidental encoding and show that the TCE is not due to task-specific strategies implemented by controlled encoding process (cf. Hintzman, 2016).

But these results also point to a large gap in our understanding of how encoding processes influence contiguity. First, many models of the TCE do not implement encoding control processes (but see Lehman & Malmberg, 2013) and would likely have difficulty accounting for the difference between the Explicit and Incidental conditions. Second, although models have simulated different processing tasks by assigning them distinct contextual states (Polyn et al., 2009), it is unlikely that this would allow a model to capture the influence of type of processing on the size of the TCE. Some mechanism to allow task demands to modulate parameters such as the rate of mental context drift seems necessary.

References

- Aristotle. (n.d.). *De memoria et reminiscencia*. , 451b10.
- Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 85–121). New York: John Wiley and Sons.
- Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(4), 1179.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing : A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- Damerau, F. J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 7(3), 171–176.
- 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. doi: 10.1037/0033-295X.112.1.3
- Davis, O. C., Geller, A. S., Rizzuto, D. S., & Kahana, M. J. (2008). Temporal associative processes revealed by intrusions in paired-associate recall. *Psychonomic Bulletin & Review*, 15(1), 64–69.
- Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, 81(5), 1179–1189.
- Healey, M. K., Crutchley, P., & Kahana, M. J. (2014). Individual differences in memory search and their relation to intelligence. *Journal of Experimental Psychology: General*, 143(4), 1553–1569. doi: 10.1037/a0036306

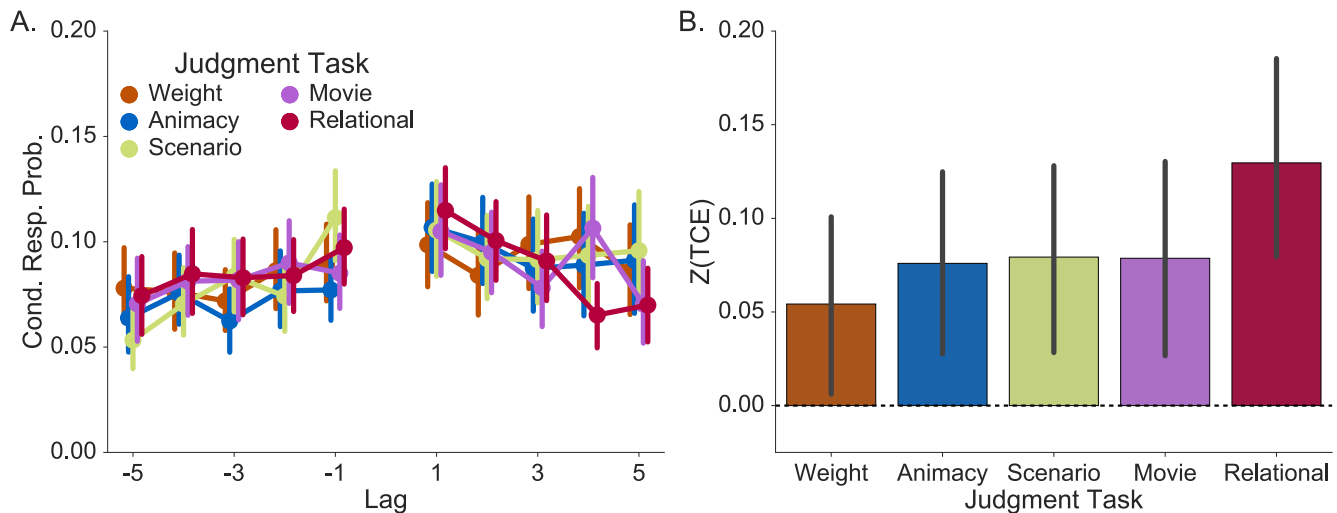


Figure 3. The temporal contiguity effect (TCE) under incidental encoding with different judgment tasks. (A) Lag-conditional response probability functions. Error bars are bootstrapped within-subject 95% confidence intervals. (B) The average $Z(\text{TCE})$. Error bars are bootstrapped between-subject 95% confidence intervals. $Z(\text{TCE})$ for a given subject is computed as follows: An observed temporal factor score was computed as the average percentile ranking the temporal lag of each actual transition in the recall sequence with respect to the lags of all transitions that were possible at that time. To determine the temporal factor score expected by chance, a permutation distribution was created by randomly shuffling the order of recalls within the sequence 10,000 times and computing a temporal factor score for each shuffling. The reported value, $Z(\text{TCE})$, is z-score of the observed temporal factor score within the permutation distribution.

- Healey, M. K., & Kahana, M. J. (2016). A four-component model of age-related memory change. *Psychological Review*, 123(1), 23-69.
- Healey, M. K., Long, N. M., & Kahana, M. J. (2106). Contiguity in episodic memory. *article submitted for publication*.
- Hintzman, D. L. (2016). Is memory organized by temporal contiguity? *Memory & cognition*, 44, 365-375.
- Jarrold, C., Hall, D., Harvey, C. E., Tam, H., Towse, J. N., & Zarandi, A. L. (2015). What can we learn about immediate memory from the development of children's free recall? *The Quarterly Journal of Experimental Psychology*, 68(9), 1871-1894.
- Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, 24(1), 103-109. doi: 10.3758/BF03197276
- Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of Learning and Motivation*, 61, 237-284.
- Kliegl, O., & Bäuml, K.-H. T. (2016). Retrieval practice can insulate items against intralist interference: Evidence from the list-length effect, output interference, and retrieval-induced forgetting. *Journal of Experimental Psychology Learning Memory and Cognition*, 42, 202-214.
- Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review*, 120(1), 155-189. doi: 10.1037/a0030851
- Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Intralist and interlist effects in free recall. *Psychological Review*, 122(2), 337-363.
- Nairne, J. S., Cogdill, M., & Lehman, M. (submitted). Adaptive memory: Temporal, semantic, and rating-based clustering following survival processing.
- Nairne, J. S., & Pandeirada, J. N. S. (in press). Adaptive memory: The evolutionary significance of survival processing. *Perspectives on Psychological Science*.
- Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, 116, 129-156. doi: 10.1037/a0014420
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93-134. doi: 10.1037/0033-295X.88.2.93
- Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). List-method directed forgetting in cognitive and clinical research: a theoretical and methodological review. *Psychology of Learning and Motivation*, 59, 131 - 189.
- Sahakyan, L., & Smith, J. R. (2014). A long time ago, in a context far, far away: Retrospective time estimates and internal context change. *Journal Experimental Psychology: Learning, Memory and Cognition*, 40(1), 86-93. doi: 10.1037/a0034250
- Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, 16, 898-904. doi: 10.1111/j.1467-9280.2005.01634.x
- Sederberg, P. B., Miller, J. F., Howard, W. H., & Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition*, 38(6), 689-699. doi: 10.3758/MC.38.6.689
- Tulving, E., & Pearlstone, Z. (1966). Availability versus access-

- sibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381-391. doi: 10.1016/S0022-5371(66)80048-8
- Turk-Browne, N. B., Simon, M. G., & Sederberg, P. B. (2012). Scene representations in parahippocampal cortex depend on temporal context. *Journal of Neuroscience*, 32(21), 7202 - 7207. doi: 10.1523/JNEUROSCI.0942-12.2012
- Unsworth, N. (2008). Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. *Journal of Memory and Language*, 59, 223-236. doi: 10.1016/j.jml.2008.04.002
- Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning Memory and Cognition*, 42, 50–61.
- Wahlheim, C. N., & Huff, M. J. (2015). Age differences in the focus of retrieval: Evidence from dual-list free recall. *Psychology and Aging*, 30(4), 768.