Navigating the "chain of command": Enhanced integrative encoding through active control of study

Douglas B. Markant (dmarkant@uncc.edu)
Department of Psychological Science
University of North Carolina at Charlotte
9201 University City Blvd., Charlotte, NC 28223 USA

Abstract

A growing body of research indicates that "active learning" improves episodic memory for material experienced during study. It is less clear how active learning impacts the integration of those experiences into flexible, generalizable knowledge. This study used a novel active transitive inference task to investigate how people learn a relational hierarchy through active selection of premise pairs. Active control improved memory for studied premises as well as transitive inferences involving items that were never experienced together during study. Active learners also exhibited a systematic search preference, generating sequences of overlapping premises that may facilitate relational integration. Critically, however, advantages from active control were not universal: Only participants with higher working memory capacity benefited from the opportunity to select premise pairs during learning. These findings suggest that active control enhances integrative encoding of studied material, but only among individuals with sufficient cognitive resources.

Keywords: active learning; transitive inference; information search; integrative encoding

Introduction

How does the opportunity to control a learning experience alter subsequent memory of it? Recent research has shown that active control over learning enhances episodic memory for experienced material compared to passive observation of the same information (Markant, DuBrow, Davachi, & Gureckis, 2014; Voss, Gonsalves, Federmeier, Tranel, & Cohen, 2011). This enhancement can arise from a number of mechanisms, including improved attentional coordination, metacognitive monitoring, or enriched encoding associated with volitional control (Markant, Ruggeri, Gureckis, & Xu, 2016).

Less is known about how active control affects the integration of studied material into flexible, generalizable knowledge. Other work has revealed benefits from active information selection when learning categorical rules (Markant & Gureckis, 2014) or causal structures (Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). However, these studies have not examined a crucial question about improvements in generalization following active control of study: Do they reflect better memory for experienced information itself (which then supports generalization later on) or the formation of relational knowledge during encoding that abstracts away from that experience? Following Zeithamova, Schlichting, and Preston (2012), these alternatives can be mapped onto two types of memory formation: *elemental encoding* of stimuli or associations that are directly experienced during study, and *integrative encoding* through which disparate study episodes are bound together into a unified representation. Whereas existing research has established that active control enhances elemental encoding in a variety of contexts, its relationship to integrative encoding remains unclear.

The present study examined the effects of active control in a well-known example of relational generalization: transitive inference (TI). In TI people learn about an ordered hierarchy (e.g., A $<$ B $<$ C) by studying premises comprised of adjacent items (e.g., A $<$ B, B $<$ C). They are then tested on their memory for studied pairs (*recall trials*; e.g., A ? B) and their ability to infer relationships between items that were never experienced together (*inference trials*; e.g., A ? C).

Transitive inference is a fundamental form of reasoning and has been the subject of a wealth of past research, but has always been studied under passive conditions in which control over the study experience is absent. This study introduces a novel *active transitive inference* task in which participants choose which premises to study during learning. Based on prior evidence that active selection improves episodic memory, active selection was expected to improve recall of studied premises relative to passive study. Active control was also predicted to improve transitive inference, but this advantage might arise from two distinct mechanisms. Enhanced elemental encoding of premises should bolster retrieval at the time of test, allowing participants to make transitive inferences by reasoning across overlapping pairs. Alternatively, active control may enhance integrative encoding during study, aiding the formation of a unified representation of the hierarchy. Importantly, these processes predict distinct relationships between performance and the distance between test items (see below), making TI well-suited to examine how learner control changes the representation of studied material.

Elemental vs. integrative encoding in transitive inference

Transitive inference involves comparing items that have never been experienced together but are linked by one or more studied pairs. TI may be supported by a number of alternative processes which can be distinguished by their dependence on elemental or integrative encoding. Elemental encoding-based inference occurs by reactivating studied premises at the time of test and reasoning across overlapping relations (Kumaran & McClelland, 2012). In this case, successful inference hinges

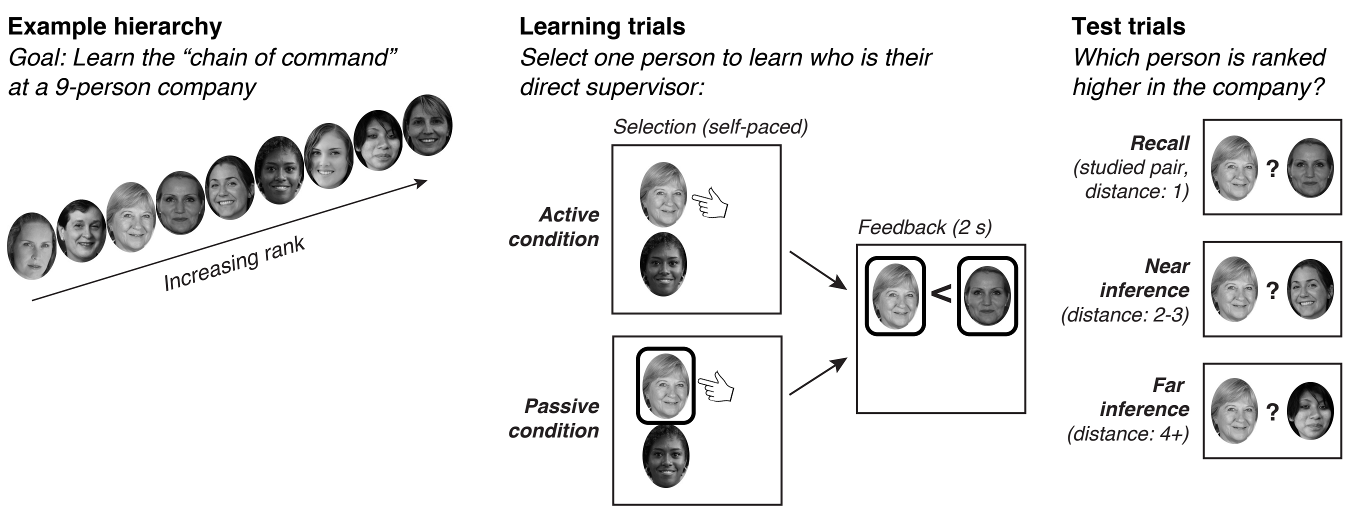

Figure 1: Depiction of the transitive inference task.

on robust encoding of studied pairs to ensure later retrieval. This process implies that large distances between test items (e.g., A ? E) will be associated with lower accuracy since there are more opportunities for retrieval errors along the way.

In contrast, integrative encoding-based accounts of TI postulate the formation of a unified, ordinal representation during study (De Soto, London, & Handel, 1965; Hummel & Holyoak, 2001; Shohamy & Wagner, 2008; Trabasso, Riley, & Wilson, 1975). Inference then entails comparing the positions of any two items along that dimension. Importantly, this process implies accuracy should *increase* with inferential distance, as items that are further apart on that latent dimension are easier to distinguish. Such *symbolic distance effects* are a hallmark of integrative encoding (Moyer & Landauer, 1967).

Although alternative forms of associative or reinforcement learning may also support TI (Frank, Rudy, Levy, & O'Reilly, 2005), the construction of an integrated representation during encoding is especially likely when participants are aware there is an underlying hierarchy to be learned (Greene, Spellman, Levy, Dusek, & Eichenbaum, 2001; Lazareva & Wasserman, 2010). Accuracy is higher among participants who report post-task awareness of the hierarchy (Martin & Alsop, 2004), who are informed about it prior to training (Greene et al., 2001; Smith & Squire, 2005), or when stimuli evoke hierarchical schemas (Kumaran, 2013). Reliance on integrated representations also appears to depend on working memory capacity (WMC) (Titone, Ditman, Holzman, Eichenbaum, & Levy, 2004; Fales et al., 2003). Thus, while constructing an integrated representation is typically associated with superior generalization, it may also depend on explicit awareness of the hierarchical organization of items and incur greater cognitive costs.

Learner control and integrative encoding

Since elemental and integrative encoding predict distinct relationships between inferential distance and performance, TI can be used to examine whether active control has broader benefits for memory formation beyond improved episodic memory for studied pairs. One reason to expect enhanced integrative encoding is that the opportunity to select premises may encourage learners to construct an integrated representation as they learn, which can then guide selection decisions (e.g., allocating study to items from less familiar portions of the hierarchy). At the same time, this process might involve additional demands on aspects of executive functioning such as working memory. To evaluate this possibility an assessment of WMC (operation span) was included in addition to the TI task in the experiment below.

In addition to the main goal of identifying any effect of active control on integrative encoding, the TI task was designed to explore information search during active study. Passive training in TI is often scaffolded such that overlapping pairs are experienced in direct succession (e.g., A $<$ B, B $<$ C, ...), which leads to faster learning than random sequences (Halford, 1984). If studying overlapping premises aids relational integration, active learners may prefer to select such options when possible. Each selection therefore involved a choice between a *near* and *far* option which differed in their distance from the pair studied on the previous trial. This made it possible to identify any search preference during active study and its relationship to inferential accuracy.

Experiment

Participants and Materials

N = 100 participants (60 women; age: M = 21.94 years, SD = 5.60) were recruited from the student population at UNC Charlotte. Participants received either course credit or $8 ($4 per session), as well as a $0–$5 incentive based on their performance in the first test session. N = 62 participants returned for the second session.

Face stimuli for the TI task were obtained from the 10k US Adult Faces Database (Bainbridge, Isola, & Oliva, 2013). For

Table 1: Estimated fixed effects from mixed effects logistic regression model of test accuracy.

Predictor	OR	95% CI-lower	95% CI-upper	Wald z	p
(Intercept)	4.07	3.30	5.07	12.36	0.00
Condition [passive]	0.76	0.67	0.85	-5.19	0.00
Session [retest]	0.85	0.76	0.98	-2.67	0.01
Distance	1.10	1.02	1.16	3.00	0.00
Operation span	2.06	1.67	2.58	6.47	0.00
Condition [passive] x Session [retest]	0.80	0.68	0.94	-2.68	0.01
Condition [passive] x Distance	0.94	0.87	1.02	-1.33	0.18
Condition [passive] x Operation span	0.55	0.49	0.60	-12.49	0.00

each sex, the stimulus set was filtered to include only faces that were non-famous and which had mean ratings within a 1-point interval centered on the midpoint of the rating scale for perceived age, emotional affect, and memorability. Thirty-six images (18 male, 18 female) were manually chosen from the filtered set to ensure high image quality and the absence of other distinctive features (e.g., jewelry, background objects).

Procedure

There were two sessions. The first session included the TI task followed by the operation span assessment. The second session occurred 6-8 days later and included a second run of the test phases from the TI task.

The TI task (Figure 1) used a within-subjects design with two rounds. Participants were tasked with learning the "chain of command" at two companies. Each participant learned about one 9-item hierarchy in the active condition and a second 9-item hierarchy in the passive condition. Each hierarchy was composed of all female faces or all male faces in order to reduce interference between conditions. The order of conditions and mapping of stimulus set to condition were counterbalanced across participants. Each round was comprised of a learning phase (56 trials) followed by a test phase (72 trials).

The instructions included an example of a 3-item hierarchy in which participants learned about two premise pairs (person A $<$ person B, person B $<$ person C) and were asked to infer the transitive relation (person A $<$ person C). All participants were therefore aware of the hierarchical nature of the stimuli and were explicitly instructed to learn to judge the relative rank of any two individuals in a given company.

Learning phase. The learning phase involved a series of choices between two non-adjacent items in the present hierarchy (excluding the highest-ranking item which was never presented as a choice option). The options on the first learning trial were two non-adjacent items sampled at random. On all subsequent trials, options differed in their distance from the item selected on the previous trial: Each option set included a *near* option that was 1–2 positions away from the item selected on the previous trial, and a *far* option that was 3 or more positions away from the item selected on the previous trial. This manipulation of option distance was designed to test whether participants in the active condition preferred to select items based on their distance. In the passive condition selections were evenly divided between near and far options.

Active study condition. Each trial began with the presentation of the two options in a vertical array in random order (Figure 1, middle). Participants were instructed to select an option at their own pace in order to learn that person's direct supervisor. Following their choice the unselected option disappeared and the premise pair (selected item and superordinate feedback item) was displayed for 2 s.

Passive study condition. In the passive condition participants did not decide which option to select. As in the active condition, the trial began with the presentation of two options, one of which was already highlighted with a red border. Participants were instructed to select the highlighted option at their own pace, at which point the trial proceeded in the same manner as in the active condition.

Test phase. In each test trial, two items were presented side-by-side and the participant clicked on the person they judged to be ranked higher in the hierarchy. The test phase was comprised of three trial types (Figure 1, right): *recall* trials involving a choice between studied premise pairs (e.g., A ? B), *near inference* trials involving items that were 2–3 positions apart (e.g., A ? C), and *far inference* trials involving items that were 4 or more positions apart (e.g., A ? E). In the second session, participants completed a second run of the same test phases experienced during the first session, with test pairs presented in a new random order.

Operation span. In the operation span task, participants attempt to hold a sequence of items in memory while judging the validity of interleaved math operations (Unsworth, Heitz, Schrock, & Engle, 2005). At the end of a trial involving multiple such steps, participants recall the sequence of digits in the same order as they appeared. The set size (number of operations/digits) ranged from 2–7, presented in increasing order, with three trials completed for each set size. Participants were highly accurate at evaluating the validity of the math operations (judgment accuracy M = 0.92, SD = 0.06). Operation span was scored according to the summed number of digits recalled in the correct order for those trials in which no

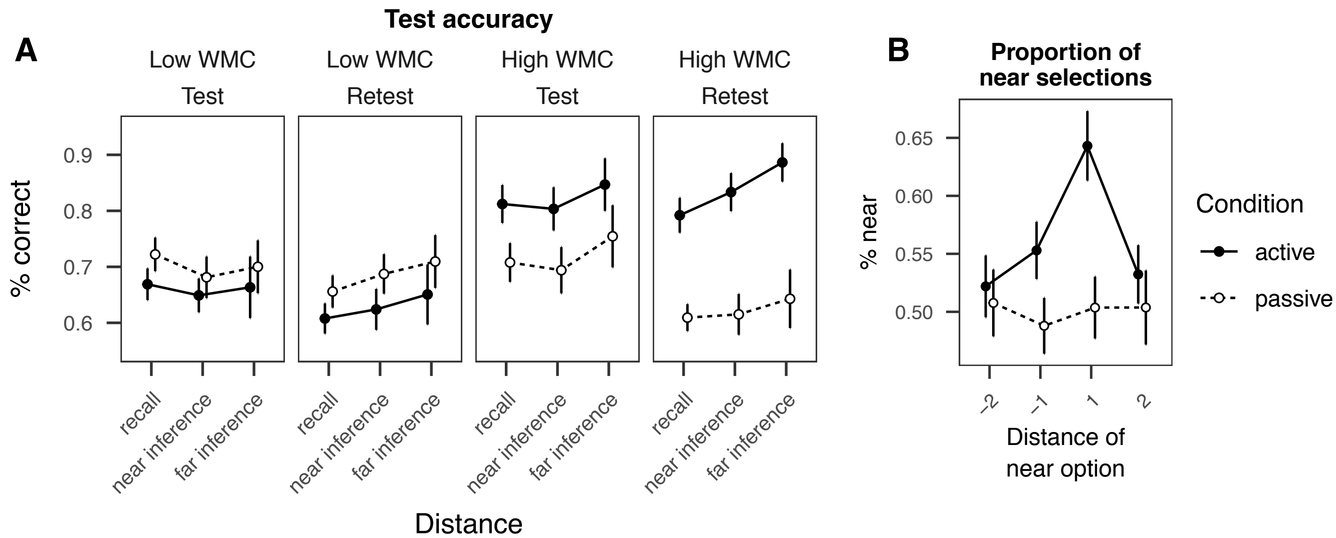

Figure 2: **A:** Test accuracy after median split on operation span, for both the immediate test and delayed retest. Performance is shown as a function of study condition (active/passive) and inferential distance (recall/near inference/far inference). **B:** Proportion of trials in which the *near* option was selected as a function of its distance from the item chosen on the previous trial. All error bars indicate within-subjects 95% confidence intervals.

errors were made (M = 14.86, SD = 11.29, median = 12.50).

Results

Test accuracy. Test responses were scored according to whether participants correctly identified the superordinate item in each test pair (0 = incorrect, 1 = correct). Test trials involving either endpoint of the hierarchy were excluded from analysis since participants could rely on non-transitive strategies to respond. Accuracy was modeled using mixed effects logistic regression (Table 1). The model included fixed effects for condition (active/passive), session (test/retest), distance (recall/near inference/far inference), and operation span (continuous), as well as pairwise interactions between condition and the other predictors. Random intercepts were included for participants and stimuli in each test pair.

Active performance was higher than passive performance in both the immediate test (active: M = 0.74, SD = 0.21; passive: M = 0.71, SD = 0.21; OR = 1.32, CI = [1.13, 1.54], z = 5.19, $p < .001$) and in the retest (active: M = 0.73, SD = 0.21; passive: M = 0.65, SD = 0.21; OR = 1.64, CI = [1.37, 1.98], z = 7.64, $p < .001$). Accuracy declined from the immediate test to the retest in both the active condition (OR = 0.85, CI = [0.71, 1.01], z = -2.67, p = 0.007) and the passive condition (OR = 0.68, CI = [0.57, 0.81], z = -6.48, $p < .001$).

There was a symbolic distance effect, such that accuracy increased with inferential distance, in the active condition (OR = 1.10, CI = [1.00, 1.20], z = 3.00, p = 0.003) but not the passive condition (OR = 1.04, CI = [0.95, 1.13], z = 1.22, p = 0.22). In addition, operation span was positively related to accuracy in the active condition (OR = 2.06, CI = [1.50, 2.84], z = 6.47, $p < .001$) but not the passive condition (OR = 1.13, CI = [0.82, 1.54], z = 1.09, p = 0.28). Figure 2A shows test accuracy in each condition following a median split on operation span. Active control of study had markedly different consequences depending on participants' operation span, with active control leading to a large, persistent advantage over passive study only among higher WMC participants.

Selections during learning. The next analysis examined participants' selections during learning and whether they could account for differences in test performance described above. Study condition was not related to item selection frequency (multinomial logistic regression, likelihood ratio test: $\chi^2_{(1,7)}$ = 7.20, p = 0.41), indicating that the aggregate distribution of experienced premise pairs was comparable across active and passive study.

Each learning trial involved a choice between a near option (1–2 positions away from the option selected on the previous trial) and a far option (3+ positions away). By design, near and far options were chosen with equal frequency during passive study. In the active condition participants had a small but significant preference for selecting the near option (M = 0.56, SD = 0.07; OR = 1.30, CI = [1.21, 1.40], z = 6.86, $p < .001$).

Near selections may be especially useful if they cause overlapping premise pairs to be experienced in successive trials, which could facilitate integrative encoding when representations of overlapping premise pairs are simultaneously active. I next examined whether the preference to select near items depended on the distance between the near option and the item selected on the previous trial ($dist_{\text{near}} \in \{-2,-1,+1,+2\}$). When $dist_{\text{near}} = +1$, the near option was immediately superordinate to the previously selected item; that is, the near option had appeared as the feedback in the previous trial.

Figure 2B shows the proportion of near selections as a function of near option distance. In the active condition, the proportion of near selections did not differ from the passive condition when $dist_{\text{near}} = -2$ (*OR* = 1.05, *CI* = [0.86, 1.29], z = 0.63, p = 0.53) or $dist_{\text{near}} = +2$ (*OR* = 1.15, *CI* = [0.94, 1.41], z = 1.73, p = 0.08). However, there was a higher proportion of near selections when $dist_{\text{near}} = -1$ (*OR* = 1.29, *CI* = [1.07, 1.56], z = 3.48, $p < .001$) or $dist_{\text{near}} = +1$ (*OR* = 1.75, *CI* = [1.45, 2.11], z = 7.52, $p < .001$). Within the active condition, the proportion of near selections was markedly higher for $dist_{\text{near}} = +1$ than $dist_{\text{near}} = -1$ options (*OR* = 1.45, *CI* = [1.20, 1.75], z = 4.95, $p < .001$). In the active condition participants therefore preferred the near option when it was adjacent to the item selected on the previous trial, and this preference was strongest when the option had appeared as feedback in that trial. Although the aggregate frequency of item selection was similar across conditions, this result suggests that active participants generated study sequences in which overlapping premise pairs were more likely to be experienced in successive trials.

Can this tendency to select overlapping items account for the performance benefit in the active condition? A new model of test accuracy was fit for the active condition which included predictors for the proportion of near selections at each level of $dist_{\text{near}}$. There were no significant relationships between accuracy and the proportion of near selections at any distance ($dist_{\text{near}} = -2$: *OR* = 1.01, *CI* = [0.70, 1.44], z = 0.06, p = 0.95; $dist_{\text{near}} = -1$: *OR* = 1.30, *CI* = [0.92, 1.85], z = 1.87, p = 0.06); $dist_{\text{near}} = +1$: *OR* = 0.96, *CI* = [0.67, 1.38], z = -0.27, p = 0.79; $dist_{\text{near}} = +2$: *OR* = 1.30, *CI* = [0.90, 1.89], z = 1.75, p = 0.08). The proportion of near selections at any distance was also unrelated to operation span ($dist_{\text{near}} = -2$: *OR* = 0.93, *CI* = [0.80, 1.07], z = -1.34, p = 0.18; $dist_{\text{near}} = -1$: *OR* = 0.98, *CI* = [0.85, 1.12], z = -0.38, p = 0.70; $dist_{\text{near}} = +1$: *OR* = 1.07, *CI* = [0.93, 1.22], z = 1.19, p = 0.23; $dist_{\text{near}} = +2$: *OR* = 0.98, *CI* = [0.85, 1.14], z = -0.30, p = 0.77). Thus, the preference to select overlapping options was a general one and could not on its own account for the gap between active and passive performance.

Discussion

This study used a novel TI task to examine whether active control aids the integration of relational knowledge during study. Control over the selection of premise pairs improved performance relative to passive study in both an immediate test and a retest one week later. Symbolic distance effects observed in the active condition strongly imply that this benefit resulted from enhanced integrative encoding, such that active learners relied on an integrated representation of the hierarchy rather than sequential reactivation of premise pairs at test (Acuna, Sanes, & Donoghue, 2002; Zeithamova, Schlichting, & Preston, 2012). The absence of such effects following passive study suggests that integrative encoding was less prevalent when the same participants lacked the opportunity to select premises for themselves.

Active control did not benefit all learners, however, as working memory capacity strongly predicted accuracy in the active condition. Among higher WMC participants, active control produced a $\sim$10% initial advantage over passive study (increasing to $\sim$20% in the retest) and sustained performance across sessions. WMC was unrelated to accuracy in the passive condition, a finding that conflicts with reports that WMC moderates TI under experimenter-controlled conditions (Fales et al., 2003; Libben & Titone, 2008; Titone et al., 2004). This discrepancy may be due to the relative difficulty of passive study in the present task. Previous studies have typically involved smaller hierarchies and scaffolded training sequences in which participants are likely to experience overlapping premises (e.g., Libben & Titone, 2008). With larger hierarchies and greater distances between successive premises, the passive condition used here may have been especially difficult even for participants with higher WMC. An important next step is to evaluate whether the large disadvantage from passive study among higher WMC persists when observing more useful sequences of premises (e.g., when yoked to participants' selections in the active condition).

This study provides the first evidence of systematic search in active TI: Participants strongly preferred to select options that appeared as feedback on the previous trial ($dist_{\text{near}} = +1$). They thereby naturally generated "chained" sequences of overlapping pairs which tend to improve performance in passive conditions relative to random presentation (Halford, 1984). This preference was widespread: 73 of 100 participants chose the $dist_{\text{near}} = +1$ option in more than half of trials in which one appeared, and the proportion of near selections was unrelated to WMC. Although selection of overlapping pairs should facilitate integrative encoding, not everyone benefited from it. One possibility is that only higher WMC individuals capitalize on chained sequences because they maintain representations of premises from trial to trial. Alternatively, higher WMC individuals may be more likely to use an integrated representation of the hierarchy to decide which option to study next (e.g., choosing to learn about the option whose rank is more uncertain). Further work is necessary to determine whether this goal-directed evaluation of options' usefulness during selection contributes to the active advantage among higher WMC individuals.

Finally, it is important to note that participants in this study were aware that there was an underlying hierarchy to be learned. Awareness influences strategy use in TI (Smith & Squire, 2005) and it is unknown how active control might affect performance in its absence. It is likely that active control would enhance elemental encoding in such conditions, perhaps due to the mere opportunity for volitional control (Murty, DuBrow, & Davachi, 2015) or additional metacognitive processing (Kornell, Klein, & Rawson, 2015). An intriguing further possibility is that active control increases the likelihood of becoming aware of an underlying hierarchy by focusing attention on abstract relationships across study episodes (Henriksson & Enkvist, 2016). This would lend support to the broader notion that active learning not only enriches memory for experienced materials, but also fosters self-directed discovery of abstract, relational knowledge.

References

Acuna, B. D., Sanes, J. N., & Donoghue, J. P. (2002). Cognitive mechanisms of transitive inference. *Experimental Brain Research*, *146*(1), 1–10.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323.

De Soto, C. B., London, M., & Handel, S. (1965). Social reasoning and spatial paralogic. *Journal of Personality and Social Psychology*, *2*(4), 513.

Fales, C. L., Knowlton, B. J., Holyoak, K. J., Geschwind, D. H., Swerdloff, R. S., & Gonzalo, I. G. (2003). Working memory and relational reasoning in Klinefelter syndrome. *Journal of the International Neuropsychological Society*, *9*(6), 839–846.

Frank, M. J., Rudy, J. W., Levy, W. B., & O'Reilly, R. C. (2005). When logic fails: Implicit transitive inference in humans. *Memory & Cognition*, *33*(4), 742–750.

Greene, A. J., Spellman, B. A., Levy, W. B., Dusek, J. A., & Eichenbaum, H. B. (2001). Relational learning with and without awareness: Transitive inference using nonverbal stimuli in humans. *Memory & cognition*, *29*(6), 893–902.

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology*, *16*(1), 65–93.

Henriksson, M. P., & Enkvist, T. (2016). Learning from observation, feedback, and intervention in linear and nonlinear task environments. *The Quarterly Journal of Experimental Psychology*, 1–57.

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In *Spatial schemas in abstract thought* (pp. 279–305).

Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 283.

Kumaran, D. (2013). Schema-driven facilitation of new hierarchy learning in the transitive inference paradigm. *Learning & Memory*, *20*(7), 388–394.

Kumaran, D., & McClelland, J. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573.

Lazareva, O. F., & Wasserman, E. A. (2010). Nonverbal transitive inference: Effects of task and awareness on human performance. *Behavioural Processes*, *83*(1), 99–112.

Libben, M., & Titone, D. (2008). The role of awareness and working memory in human transitive inference. *Behavioural processes*, *77*(1), 43–54.

Markant, D., DuBrow, S., Davachi, L., & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory & Cognition*, *42*(8), 1211–1224.

Markant, D., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94–122.

Markant, D., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, *10*(3), 142–152.

Martin, N., & Alsop, B. (2004). Transitive inference and awareness in humans. *Behavioural processes*, *67*(2), 157–165.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.

Murty, V. P., DuBrow, S., & Davachi, L. (2015). The simple act of choosing influences declarative memory. *The Journal of Neuroscience*, *35*(16), 6255–6264.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*(2), 378–389.

Smith, C., & Squire, L. R. (2005). Declarative memory, awareness, and transitive inference. *Journal of Neuroscience*, *25*(44), 10138–10146.

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2011). Connecting instances to promote children's relational reasoning. *Journal of experimental child psychology*, *108*(2), 260–277.

Steyvers, M., Tenenbaum, J., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*(3), 453–489.

Titone, D., Ditman, T., Holzman, P. S., Eichenbaum, H., & Levy, D. L. (2004). Transitive inference in schizophrenia: impairments in relational memory organization. *Schizophrenia research*, *68*(2-3), 235–247.

Trabasso, T., Riley, C. A., & Wilson, E. (1975). The representation of linear order and spatial strategies in reasoning: A developmental study. *Reasoning: Representation and process in children and adults*, 201–229.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods*, *37*(3), 498–505.

Voss, J., Gonsalves, B., Federmeier, K., Tranel, D., & Cohen, N. (2011). Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. *Nature Neuroscience*, *14*(1), 115–120.

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in Human Neuroscience*, *6*.