The role of task understanding on younger and older adults' performance

By: David J. Frank and Dayna R. Touron

Frank, D. J. & Touron, D. R. (2019). The role of task understanding on younger and older adults' performance. *Journal of Gerontology: Psychological Sciences, 74(2),* 264-274.

This is a pre-copyedited, author-produced version of an article accepted for publication in *Journal of Gerontology: Psychological Sciences* following peer review. The version of record, David J Frank, Dayna R Touron, "The Role of Task Understanding on Younger and Older Adults' Performance," *The Journals of Gerontology: Series B*, Volume 74, Issue 2, February 2019, Pages 264–274, is available online at: https://doi.org/10.1093/geronb/gbw165

Abstract:

Objectives: Age-related performance decrements have been linked to inferior strategic choices. Strategy selection models argue that accurate task representations are necessary for choosing appropriate strategies. But no studies to date have compared task representations in younger and older adults. Metacognition research suggests age-related deficits in updating and utilizing strategy knowledge, but other research suggests age-related sparing when information can be consolidated into a coherent mental model.

Method: Study 1 validated the use of concept mapping as a tool for measuring task representation accuracy. Study 2 measured task representations before and after a complex strategic task to test for age-related decrements in task representation formation and updating.

Results: Task representation accuracy and task performance were equivalent across age groups. Better task representations were related to better performance. However, task representation scores remained fairly stable over the task with minimal evidence of updating.

Discussion: Our findings mirror those in the mental model literature suggesting age-related sparing of strategy use when information can be integrated into a coherent mental model. Future research should manipulate the presence of a unifying context to better evaluate this hypothesis.

Keywords: Concept mapping | Mental models | Strategies

Article:

Imagine you are playing a new strategy game. How will you do? Will you perform better the second time you play the game? You probably expect to do better the second time because you will have a better sense of how the game works and what strategies are effective. That is, you expect to have a better understanding of the rules, structure, goals, and mechanics of the task—a better task representation. Imagine that you were playing this game with your parents or grandparents. How will they do? If you are uncertain how aging would affect the ability to form, utilize, and update task representations, then you are not alone. Little research has focused on the role of aging and task representations on strategic choices.

Older adults often choose less effective strategies compared with younger adults in cognitive tasks, potentially contributing to their poorer performance (Brigham & Pressley, 1988; Lemaire, 2010; Price, Hertzog, & Dunlosky, 2008; Touron & Hertzog, 2004a, 2004b). Whereas younger adults adjust their strategies following task experience, older adults are more likely to repeat ineffective strategies (Brigham & Pressley, 1988; Lemaire, 2010; Price et al., 2008; Touron & Hertzog, 2004a, 2004b). Older adults' continued use of ineffective strategies has been linked to inaccurate beliefs (Price et al., 2008), poor confidence (Touron & Hertzog, 2004a,b), and habit (Lemaire, 2010). Additionally, it is possible that older adults do not match their strategies to fit the task because they have an inaccurate representation, or mental model, of the task itself. Models of strategy selection suggest that accurate task representations are critical for effective strategy selection (Bromme, Pieschl, & Stahl, 2009; Lovett & Schunn, 1999; Muis, 2007; Winne & Hadwin, 1998). For example, when playing the game Scrabble, one can earn additional points by creating more than one word, as when using an "s" in their new word to pluralize a word from an earlier turn. Models of strategic choice also assume that people identify and correct inaccurate task representations and that these corrections lead to more appropriate strategic choices later in the task. To our knowledge, no studies to date have directly tested this specific assumption.

The next section briefly describes the role of metacognition in task representation updating and research on strategy knowledge updating suggesting that older adults' ability to update and utilize knowledge may be impaired. The subsequent section describes findings from the mental model literature suggesting that when information can be represented as a cohesive mental unit, older adults demonstrate updating and utilization similar to younger adults. Thus, these literatures make opposite predictions regarding aging and task representation updating. Lastly, we describe two studies that measure task representation accuracy in younger and older adults to test whether aging negatively affects task representation formation, updating, and utilization.

Metacognition and Knowledge Updating

In order to update a faulty task representation, one must monitor the task and compare it with one's own task representation, identifying incongruencies. This process is broadly known as metacognitive monitoring (Nelson & Narens, 1990) and is generally spared with age (Dunlosky, Baker, Rawson, & Hertzog, 2006; Dunlosky & Connor, 1997; Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Hertzog, Sinclair, & Dunlosky, 2010; Kuhlmann & Touron, 2011). That is, older adults are typically as accurate as younger adults when monitoring their performance on a given trial. However, older adults often struggle to update their knowledge of strategy effectiveness or alter their strategic choices—processes known as metacognitive control (Dunlosky et al., 2006; Dunlosky & Connor, 1997; Dunlosky & Hertzog, 2000; Hertzog, Dunlosky, & Sinclair, 2010; Hertzog et al., 2002; Hertzog, Sinclair, & Dunlosky, 2010; Matvey, Dunlosky, Shaw, Parks, & Hertzog, 2002). Likewise, in causal learning tasks and probabilistic learning tasks—where participants learn to predict an outcome based on a series of cues—older adults struggle to learn about negative cue-probability relationships and use fewer cues when making predictive judgments (Chasseigne et al., 2004; Mata, von Helversen, & Rieskamp, 2010; Mutter & Asriel, 2016; but see Hines, Hertzog, & Touron, 2015). These results suggest that, relative to younger adults, older adults may also struggle to update their task representations or may not fully take advantage of accurate task representations (a utilization failure).

Aging and Mental Models

In contrast to older adults' struggles to update and utilize knowledge about strategies, older adults often demonstrate age-related sparing when information can be represented as a coherent mental model (Castel, 2007; Gilbert, Rogers, & Samuelson, 2004; Morrow, Leirer, Altieri, & Fitzsimmons, 1994a; Radvansky, 1999b; Radvansky, Copeland, Berish, & Dijkstra, 2003; Radvansky & Dijkstra, 2007; Radvansky, Gerard, Zacks, & Hasher, 1990; Radvansky, Zacks, & Hasher, 1996; Stine-Morrow, Gagne, Morrow, & DeWall, 2004; Stine-Morrow, Morrow, & Leno, 2002). For example, older adults are as adept as younger adults at updating the motives and locations of characters in a narrative (Gilbert et al., 2004; Morrow et al., 1994a; Radvansky, 1999b; Radvansky et al., 2003; Stine-Morrow et al., 2004, 2002). Older adults also demonstrate superior memory when information can be consolidated into a single mental model. For example, when given a list of objects, each in a different location, older adults perform poorly at recalling the individual objects (Radvansky et al., 1996). However, when given a list of objects located in a single location—which allows for a single integrated mental model—age differences in object memory are reduced. Thus, the mental model literature suggests that if older adults organize their task representations into cohesive units, they should show age-related sparing in mental model formation, updating, and utilization.

We used a concept mapping technique to measure task representation accuracy in younger and older adults. Concept mapping has been used extensively to measure young adults' mental models in domains ranging from accounting (Curtis & Davis, 2003), to radar warning systems (Rowe & Cooke, 1995), to negotiations (Van Boven & Thompson, 2003). However, no studies to our knowledge have used concept mapping to measure older adults' mental models. Thus, Study 1 tested whether concept mapping can measure older adults' task representations. Study 2 tested whether task representations and task representation updating influence performance on a strategic choice task—a claim prevalent in models of strategic choice. Additionally, Study 2 tested whether older adults struggle to form, update, and utilize task representations—as suggested by studies of knowledge updating—or whether they show age-related sparing—as suggested by studies of mental models.

Study 1

The purpose of Study 1 was to establish that concept mapping techniques could provide an accurate measure of task representation in younger and older adults. To minimize the influence of prior knowledge and beliefs on task representation formation, we created a novel chemistry task with a well-defined structure (analogous to managing swimming pool chemistry). To avoid the possibility of floor performance in older adults, we used guided learning quizzes and a cumulative criterion test during the task instructions. Although the criterion test required a minimum level of "declarative knowledge," previous studies show declarative knowledge is necessary but not sufficient for developing a coherent mental model (Goldsmith, Johnson, & Acton, 1991; Novak, 1990; Novak & Cañas, 2006; Staggers & Norico, 1993). Thus, variability in task representations was expected despite this learning criterion.

Critically, Study 1 used an inference test (e.g., Goldsmith et al., 1991; Novak, 1990; Novak & Cañas, 2006) to establish whether we could measure task understanding. After establishing the validity and viability of our task representation measure we could test whether task representations affect performance on the actual chemistry task (Study 2).

Methods

Participants

Twenty younger adults (aged 18–22 years) and 28 older adults (aged 61–77 years) participated. Younger adults received partial course credit whereas older adults were paid for participation. Participants were in relatively good health and none reported having ever suffered a major seizure or stroke. All participants had corrected near visual acuity of 20/50 or better. Participant demographics can be found in Table 1.

Table 1. Study 1 Means and Standard Deviations for Demographic and Performance Variables

	Young		**Old**		**Age difference *d***
Demographics					
N	20		28		
Age	18.81	(1.81)	69.49	(4.32)	
Education	12.51	(0.86)	15.49	(2.48)	+1.61**
Medications	0.57	(0.90)	2.51	(1.64)	+1.47**
Processing speed	38.62	(7.81)	27.80	(4.89)	−1.66**
Vocabulary	13.45	(3.12)	22.75	(7.26)	+1.66**
Performance					
Cumulative test	0.85	(0.13)	0.75	(0.11)	−0.83**
C-score	0.26	(0.09)	0.29	(0.09)	+0.33
Attempts to criterion	1.67	(0.91)	2.05	(0.94)	+0.41
Inference test	0.59	(0.16)	0.56	(0.17)	−0.18

Notes: Young = young adult means; Old = older adult means; Processing speed = number correct on out of 30 on the Salthouse pattern comparison task (1993); Vocabulary = number correct out of 36 on the Advanced Vocabulary Test (Ekstrom, French, & Harman, 1976); Cumulative test = proportion correct on a 15-question, 9-choice, multiple choice cumulative instructions test measuring declarative knowledge of the instructions; C-score = proportion of concept map overlap between the participant and reference models ranging from 0 (*no overlap*) to 1 (*perfect overlap*); Attempts to reach criterion = number of attempts taken to pass the cumulative instructions test; Inference test = proportion correct on a 19-question, 4-choice, multiple choice test inference test measuring the ability to apply the information from the instructions. Values in parentheses indicate standard deviations. Age difference *d* = Cohen's *d* for age comparisons.
*$p < .05$. **$p < .01$.

Materials and procedures

The study uses a novel chemistry task with a structure analogous to the daily management of swimming pool temperature, chlorine, pH, and alkalinity. However, the task is disguised as maintaining a chemical solution in a beaker. Terms normally associated with a swimming pool are replaced by lettered names (e.g., Chemical C; Property Q) except for temperature and the pool heater (the latter was renamed the Bunsen burner). The swimming pool task was chosen because properties interact with each other with varying complexity. All materials were programmed in E-Prime 2 (Schneider, Eschman, & Zuccolotto, 2002) unless otherwise stated.

Demographics

Participants first completed a demographics questionnaire, a near visual acuity test, pattern comparison test, and a vocabulary test (Table 1).

Novel chemistry task instructions

Participants next read the instructions for the chemistry task (see Supplementary Material). The instructions detailed how each chemical additive affects each chemical property and how the chemical properties interact. To aid learning, six computerized guided learning quizzes were inserted throughout the instructions, each containing one to three items. These quizzes were "open book," as participants were free to switch between the quiz screens and the instruction screens (see Supplementary Material). For incorrect responses, participants were asked to scan the instructions for the correct answers. Importantly, the instructions framed the task as managing a chemical solution in a beaker and never mentioned that it was analogous to managing a swimming pool.

After obtaining 100% accuracy on a quiz, the participant continued to the next section of the instructions. There were 15 questions, 7 relating to the properties analogous to chlorine and 8 relating to the properties analogous to pH/alkalinity.

Cumulative instructions test

After passing all mini-quizzes the participant completed a cumulative instructions test containing the same 15 items from the mini-quizzes (presented in the same order), without referencing the instructions. To pass this test, a participant had to correctly answer at minimum four of the seven "chlorine questions" and four of the eight "pH/alkalinity questions." Participants who did not meet this criterion were returned to the start of the learning phase and asked to restudy the instructions and their quiz answers. (Initial piloting did not include the mini-quizzes and included minimal feedback on criterion test performance. Under these initial conditions most participants did not reach the criterion and all voiced considerable frustration with the task. Adding the mini-quizzes and increasing the degree of feedback improved criterion test performance, eliminated complaints of frustration, and improved compliance.)

Task representations

To measure overall task representation accuracy, we used a pathfinder-derived concept mapping technique (Goldsmith, Johnson, & Acton, 1991). A concept map includes nodes (concepts) connected by links representing the relationships among concepts (Novak & Cañas, 2006; Novak, 1990). For example, in Figure 1, "Bunsen burner," "temperature," and "bacteria" are nodes, whereas the arrows indicate links. This technique provides pairs of concepts and the participant rates how related they are, repeating this procedure for all possible pairs (Goldsmith et al., 1991). The ratings are then subjected to a pathfinder algorithm that transforms the matrix of ratings into a concept map by removing indirect relationships between nodes whenever there is a stronger more direct relationship (see Johnson, Goldsmith, & Teague, 1994 for a detailed description).

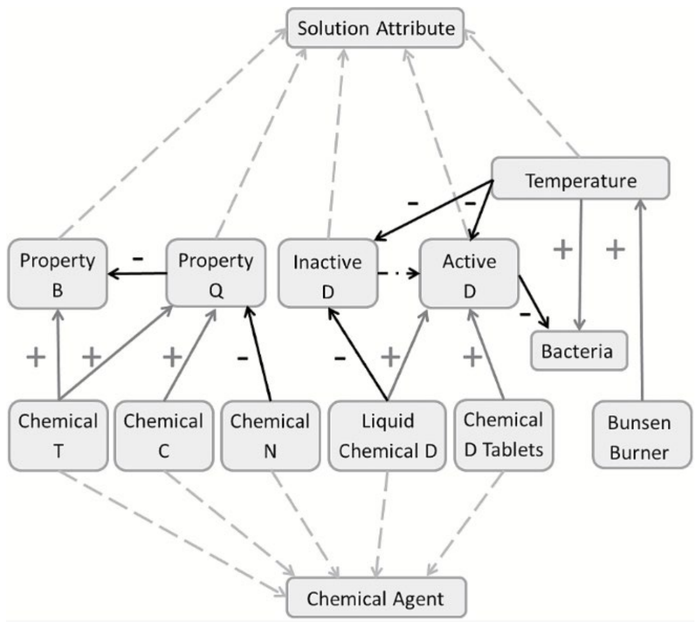

Figure 1. Study 2 reference map. Bubbles represent concepts (nodes) and links represent the relationships between nodes. Dotted light gray arrows indicate categorical relationships (i.e., Chemical C is a Chemical agent). Light gray arrows accompanied by plus (+) signs indicate positive relationships (i.e., Chemical C increases Property Q). Black arrows accompanied by minus (−) signs indicate negative relationships (i.e., Chemical N decreases Property Q). The dotted black arrow between Inactive D and Active D indicates that Inactive D is used up Active D.

The proportion of overlap between a participant's concept map and the reference map (correct concept map) is referred to as a "closeness" or "C-score" (Johnson et al., 1994). C-scores are computed as the number of shared links between a participant's concept map and the reference map, divided by the total number of links in both maps combined. Thus, C-scores range from 0 (*no overlap*) to 1 (*identical concept maps*; Goldsmith et al., 1991). C-scores from relevant concept maps predict performance in various domains, including accounting (Curtis & Davis, 2003), teaching of elementary mathematics (Gomez, Hadfield, & Housner, 1996), ACT math scores (Johnson et al., 1994), radar warning systems troubleshooting (Rowe & Cooke, 1995), and negotiations (Van Boven & Thompson, 2003). We also computed raw correlations between the participant's and an expert's pathfinder ratings and analyzed those raw correlations in place of C-scores. The results were virtually identical to those reported using C-scores.

We used Rate software (Interlink, 1990) to present 19 relevant terms (see Supplementary Material). Participants rated each possible pair of terms (presented in random order) from 1 (*unrelated*) to 9 (*related*).

Inference test

Participants then completed an inference test containing 22 multiple choice questions (see Supplementary Material). This test presented situations involving the chemistry task and asked the participants to indicate either what caused the situation or what they should do to correct it. Participants did not perform the actual chemistry task in Study 1.

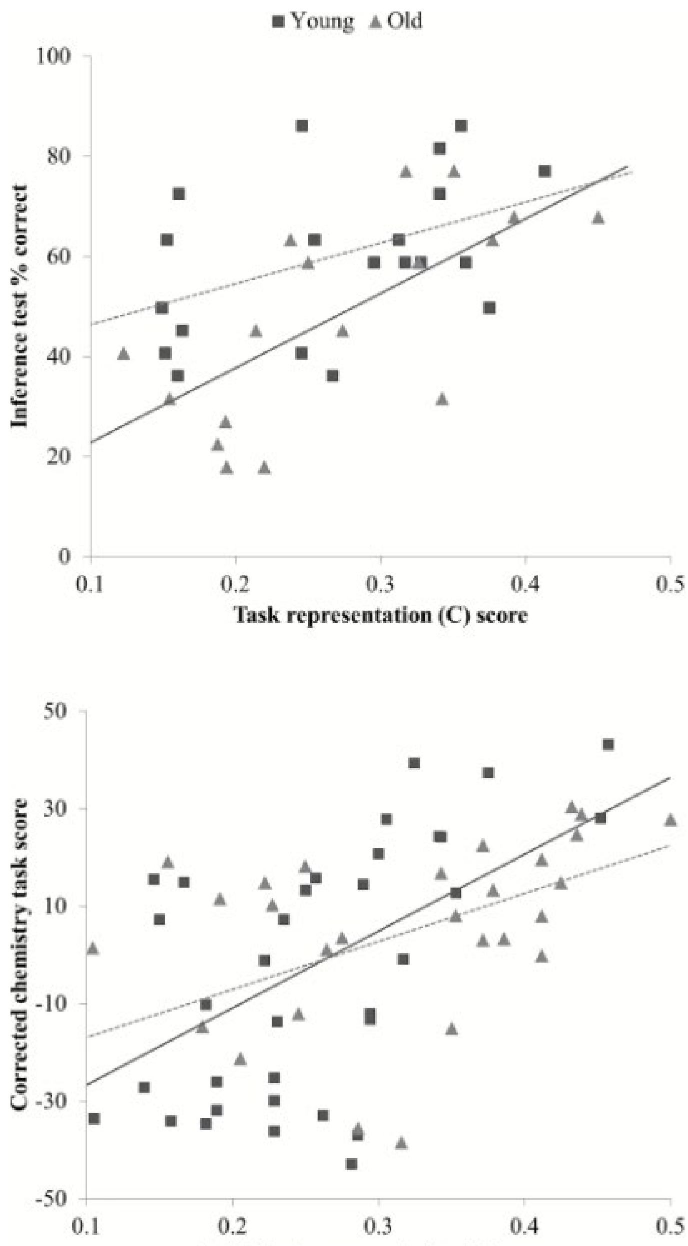

Figure 2. Relationship between C-scores and performance. Top panel: Study 1 relationship between C-scores and inference test scores (percent correct). Bottom panel: Study 2 relationship between C-scores and chemistry task performance.

Posttask measures

Participants next made a series of 1–5 ratings regarding their task experience (see Supplementary Material) and indicated any terms they felt were omitted from the ratings task (see Supplementary Material). (no participant suggested any additional terms for the ratings task). Additionally participants were asked, “What did this task make you think of?” and “Have you ever owned or managed a swimming pool before?” No participants indicated that the task reminded them of managing a swimming pool. (Participants having owned or managed a swimming pool did not perform differently on any measure relative to those not having owned or managed a pool.)

Results

Neither C-scores, nor inference test scores, differed with age (Table 1; Figure 2). These measures were positively correlated overall, $r(48) = .58$, $p = .001$, and within younger, $r(19) = .50$, and older adults, $r(27) = .67$, (Figure 2, top panel), supporting the validity of the pathfinder technique for measuring task representations in both samples.

Discussion

Despite the required level of declarative knowledge, participants’ task representation C-scores and inference test scores showed considerable variance. Higher C-scores were associated with higher test performance. This is consistent with the claim that the underlying structure of knowledge and the interrelations among ideas are more critical for knowledge utilization than is declarative knowledge (Curtis & Davis, 2003). These findings also validate the use of pathfinder-derived concept maps to measure task representations. Most importantly these results held for both age groups, validating concept mapping as a viable measure of task representations in older adults. These results provide the foundation for Study 2 which tested the impact of task representation formation, updating, and utilization on younger and older adults’ performance.

Study 2

Study 2 measured task representations before and after completing the novel chemistry task (as opposed to the inference test in Study 1) to test whether: (i) initial task representations influence performance in younger and older adults, (ii) older adults struggle to update task representations, and (iii) differences in task representation updating account for differences in performance improvements on the chemistry task. We also changed the instructions (see *Methods* section) to ensure that participants began the task with incomplete task representations that could be updated.

Methods

Participants

Thirty-five younger adults (aged 18–25 years) and 30 older adults (aged 60–75 years) participated. An additional eight older adults participated but were unable to complete the experiment in 3 hours. Participant compensation and characteristics were as described for Study 1; demographics can be found in Table 2.

Table 2. Study 2 Demographic and Performance Data

	Young		**Old**		**Age difference** ***d***
Demographics					
N	35		30		
Age	18.54	(0.85)	70.90	(3.35)	
Education	12.40	(0.69)	16.45	(1.82)	+2.94**
Medications	0.54	(0.82)	2.13	(1.98)	+1.05**
Processing speed	38.89	(7.46)	27.76	(5.34)	−1.72**
Vocabulary	15.60	(4.36)	24.70	(6.01)	+1.73**
Performance					
Cumulative test 1	0.42	(0.25)	0.47	(0.23)	+0.21
C-score 1	0.25	(0.09)	0.30	(0.12)	+0.47
Chemistry task	−2.59	(25.88)	3.01	(20.69)	+0.23
C-score 2	0.25	(0.12)	0.29	(0.07)	+0.41
Cumulative test 2	0.45	(0.31)	0.42	(0.25)	−0.11

Notes: Processing speed = number correct on out of 60 on Salthouse's Pattern Comparison task (1993); Vocabulary = number correct out of 36 on the Advanced Vocabulary Test (Ekstrom et al., 1976); Cumulative test 1 = proportion correct on a 15-question, 9-choice, multiple choice, pretask cumulative instructions test measuring declarative knowledge of the instructions; C-score 1 = pretask proportion of concept map overlap between the participant and reference models ranging from 0 (*no overlap*) to 1 (*perfect overlap*); Chemistry task = Performance on the chemistry task, positive scores indicate above average performance, negative scores indicate below average performance. C-score 2 = posttask proportion of concept map overlap; Cumulative test 2 = proportion correct on a 15-question, 9-choice, multiple choice, posttask cumulative instructions test measuring declarative knowledge of the instructions. *$p < .05$. **$p < .01$.

Materials and procedures

Participants completed the same pretests used in Study 1. (As in Study 1, participants having owned or managed a swimming pool did not perform differently on any measure relative to those not having owned or managed a pool.) Participants then completed the chemistry task instructions, ratings tasks, novel chemistry task, ratings task a second time, and posttask measures.

Novel chemistry task instructions

Pilot studies revealed that, unlike the inference test used in Study 1, younger adults performed at ceiling when performing the actual chemistry task. As a result, several changes were made to the instructions (for both age groups). The novel chemistry task instructions were presented without the guided learning quizzes. The cumulative instructions test was retained, but the criterion was eliminated and three items deemed peripheral to the task were removed (e.g., whether a property being too high caused corrosion or toxic fumes). To ensure that participants entered the novel chemistry task with imperfect task representations, we removed information regarding which chemical properties were affected by temperature and how chemicals C, T, and N impact Properties B and Q. Participants were informed that the information was incomplete and told to learn these relationships during the task.

Pretask pathfinder ratings

Participants next completed pathfinder ratings to measure their task representations following instructions. As with the cumulative instructions test, we removed items from the pathfinder ratings task that were not directly relevant to the chemistry task (see Supplementary Material). Participants were offered a break halfway through the ratings task.

Novel chemistry task

The novel chemistry task was programmed in E-Prime 2 (Schneider et al., 2002) and featured five chemical solution properties to be maintained within an ideal range (Figure 3). When properties fell outside the ideal range, the effects on the solution became more robust. Task trials were presented as a series of "days." For each day, a report indicated the current level of each property and the ideal range for each property. The participant could then add doses of chemicals or adjust the setting of the Bunsen burner. The goal of the task was to keep each solution property as close as possible to the center of its ideal range. After 7 days, a new trial began with a new chemical solution. Each solution had a different starting state, meaning different properties were outside the ideal ranges. Poor strategic choices (i.e., adding the wrong chemicals) could create further problems, whereas proper choices would eliminate problems and bring properties into their ideal ranges. However, even with ideal adjustments, some problems required multiple days to correct. Eighteen different problem starting states each occurred (randomly) once during the first half and once during the second half of the task.

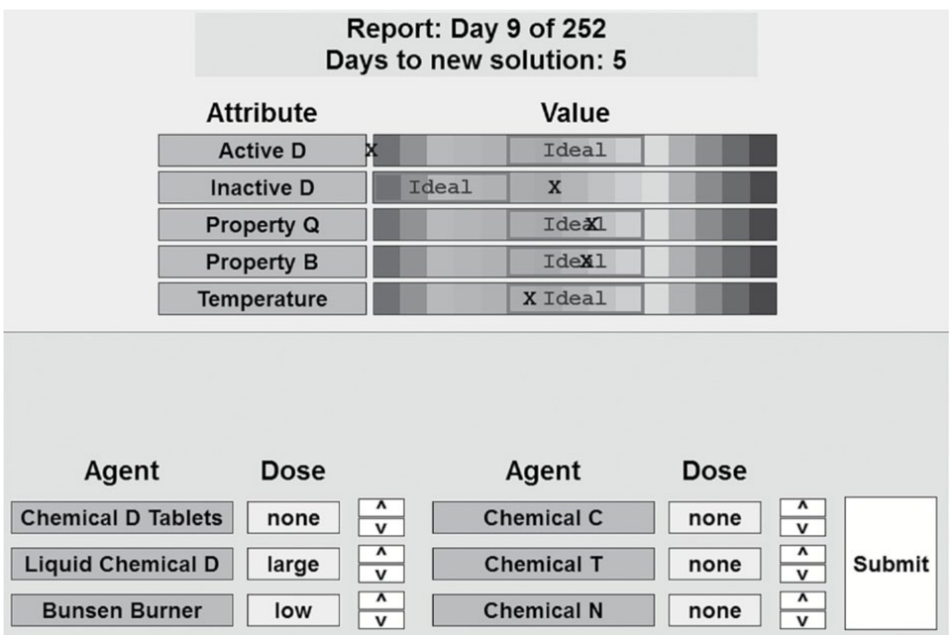

Figure 3. Chemistry task interface. The actual task was presented in color rather than grayscale.

Scoring

To control for inherent differences in starting state difficulty, scores for each solution were mean centered. Positive scores indicate above average performance (property values were closer to the midpoint) and negative scores indicate below average performance (property values were farther from the midpoint; see Supplementary Material). These scores were not visible to the participants. Instead they had to deduce the effectiveness of their strategic decisions based on whether the properties returned to their ideal ranges.

Posttask pathfinder rating and questionnaire

After the chemistry task, participants completed the ratings task again to assess changes in task representations. Finally, participants completed the instructions quiz again and the same posttask survey described in Study 1. (When asked to list any terms that they felt should have been included in the ratings task, one younger and one older adult listed link labels (verbs)—which are not typically used in pathfinder-derived concept maps. No other participants indicated that any terms were missing.)

Results

Task representations

We examined age differences in task representation formation and updating via a 2 (Age: young, old) × 2 (Time: pretask, posttask) analysis of variance. There was a trend toward younger adults entering the task with more accurate task representations, $F(1, 63) = 3.57$, $p = .063$. Neither age group demonstrated task representation updating, no main effect of time or Age × Time interaction, Fs < 1. Most participants showed only small positive or negative changes in C-scores (changes between −.1 and .1). As in Study 1, initial C-scores correlated strongly with performance $r(63) = .54$, $p < .001$, in both younger $r(32) = .54$ and older adults $r(28) = .56$ (Figure 2, bottom panel).

Novel task performance

We analyzed chemistry task scores using three hierarchical linear models (SAS Proc MIXED; Littell, Milliken, Stroup, & Wolfinger, 2000). Model 1 tested for age differences in overall performance and performance changes with practice. Model 2 tested whether initial task representations predict task performance above and beyond age and practice. Model 3 tested whether "changes" in task representations predict changes in performance over trials. For each model, younger adults served as the reference group (young = 0; old = 1), and trials were entered 0–35 so the intercept indicates average young adult performance on Trial 1.

Model 1: Changes in performance with practice

The first model testing age differences in overall performance and changes with practice (Table 3) included age, trial, and the Age × Trial interaction. The main effect of trial indicates that participants' performance on the chemistry task improved with practice. However, the null effects for age and the Age × Trial interaction suggest that younger and older adults performed similarly overall and improved similarly across trials (Figure 4).

Table 3. Model Results for Chemistry Task Performance

	Estimate	*SE*	*df*	*t* Value	*p* Value
Model 1					
Intercept	−5.65	1.80	63	−3.14	.003
Age	2.53	2.64	63	0.96	.343
Trial	0.17	0.09	2,273	1.98	.048
Age × Trial	0.18	0.13	2,273	1.35	.176
Model 2					
Intercept	−45.49	2.95	63	−15.44	<.001
Age	12.58	4.20	63	2.99	.004
Initial C-score	157.66	9.65	2,271	16.33	<.001
Trial	0.17	0.08	2,271	2.15	.032
Age × Initial C-score	−59.33	12.50	2,271	−4.75	<.001
Age × Trial	0.18	0.12	2,271	1.47	.141
Model 3					
Intercept	−50.07	2.90	63	−17.26	<.001
Age	13.75	4.86	63	2.83	.006
Initial C-score	178.71	9.59	2,267	18.64	<.001
C-score change	100.54	17.18	2,267	5.85	<.001
Trial	0.17	0.08	2,267	2.20	.028
Age × Initial C-score	−68.70	15.23	2,267	−4.51	<.001
Age × C-score change	−86.11	27.62	2,267	−3.12	.002
Age × Trial	0.18	0.12	2,267	1.53	.127
C-score change × Trial	−0.06	0.84	2,267	−0.07	.943
Age × Trial × C-score change	0.32	1.24	2,267	0.26	.798

Notes: Models of chemistry task performance. Performance scores are corrected for trial difficulty. Performance scores of zero indicate average performance, with positive values being above average and negative values being below average. Younger adults serve as the reference group. Thus positive age effects indicate better performance or steeper slopes for older adults relative to young. Main effects indicate the beta estimate for young adults with the interaction coefficient indicating how beta estimate changes for older adults.

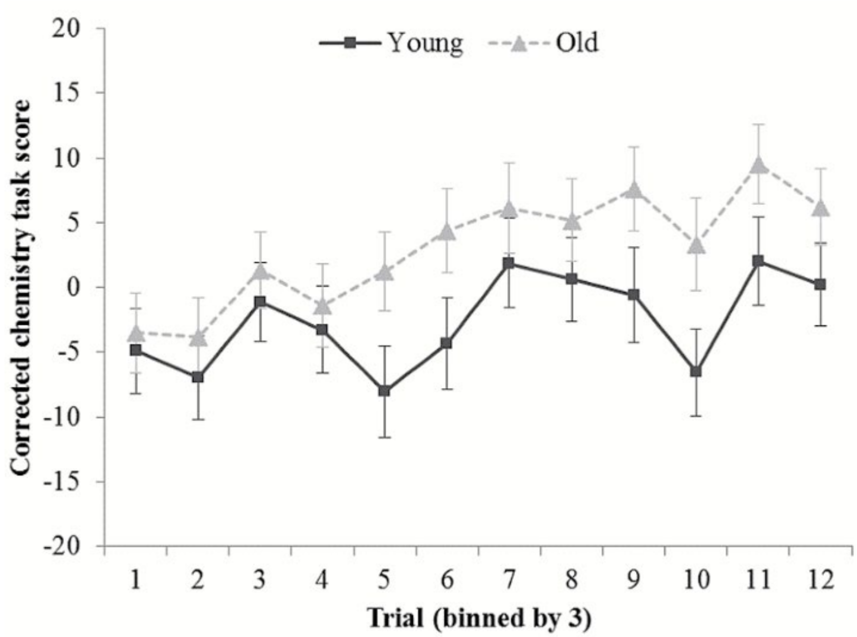

Figure 4. Study 2 Chemistry task performance over trials, by age group. Data are averaged over bins of three trials for visual purposes, but unbinned data were analyzed. Analyses of binned data produced the same pattern of results as the unbinned analyses.

Model 2: Performance as a function of initial task representation

To determine whether task representations influence performance, Model 2 added initial (pretask) C-scores and the interaction with age as predictors of chemistry task performance. The main effect of pretask C-scores confirmed that more accurate initial task representations produced better performance (Figure 2). When including pretask C-scores in the model, we obtained a significant age effect, with older adults outperforming younger adults when C-scores were held constant. However, these main effects were qualified by a significant Age × Initial C-score interaction, with a steeper slope for the initial C-score–performance relationship for younger adults than for older adults. The main effect of trial continued to be significant, and the Age × Trial interaction remained nonsignificant.

Although the main effect of age indicates higher performance in older adults, this is undermined by the Age × Initial C-score interaction. Older adults' shallower slope for initial C-scores results in higher performance estimates (relative to younger adults) when task representations are poor (initial C-score of 0.15; see Figure 2), but lower performance estimates (relative to younger adults) when task representations are even moderately accurate (initial C score = .30). In summary, accurate task representations are related to superior performance for both age groups. However, younger adults with accurate task representations may be better at utilizing them.

Model 3: Changes in performance as a function of task representation updating

Although C-scores changed little following practice, a few participants showed moderate positive or negative changes. If these changes in C-scores reflect accurate updating and forgetting/confusion respectively, this might be reflected in performance. Thus changes in C-scores and the interactions of C-score change, trial, and age were added to the model.

The main effects of initial C-scores and the Age × Initial C-score interaction remained significant, with higher initial C-scores predicting better overall task performance, and a steeper relationship for younger adults. The significant positive main effect of C-score change was qualified by a significant negative Age × C-score change interaction. This main effect indicates that improvements in C-scores from pre- to posttask were associated with higher overall performance. However, the coefficient for the age interaction indicates that, unlike younger adults, older adults showed only modest performance gains when C-scores improved.

C-scores are thought to reflect "structural" relationships between concepts. If structural changes to task representations are primarily responsible for improvements on the chemistry task, then the C-score change × Trial interaction should be significant, with greater improvements in performance for participants who corrected their task representations. However, this was not the case. The main effect of trial remained significant after accounting for C-score change. Additionally, the Change × Trial, and Age × Change × Trial interactions were not significant, indicating that the changes in task performance over trials seen in Model 1 are not reflected in changes in C-scores.

However, if changes in task representations are infrequent and occur at different points in the task for different participants, then a single linear coefficient may not describe the relationship between changes in task representations and trial-by-trial improvements. Instead, improvements resulting from task representation updating would be better fit by a main effect on performance which is germane to the specific time point during which updating occurs. This is precisely the pattern that we obtained.

Discussion

Older adults formed task representations equivalent to those of younger adults. This occurred despite removing the guided learning quizzes and cumulative test criterion. Likewise, older adults performed similar to younger adults on the novel chemistry task and similarly improved their performance. These results are consistent with studies suggesting age-related sparing in performance when information can be organized into a coherent mental model (Castel, 2007; Gilbert et al., 2004; Morrow, Leirer, Altieri, & Fitzsimmons, 1994b; Radvansky et al., 2003, 1990; Radvansky & Dijkstra, 2007; Radvansky, 1999a; Stine-Morrow et al., 2004, 2002). It is important to note that eight older adults were unable to complete Study 2 within the allotted time. Thus, the age-related sparing in our task may, in part, be a by-product of the poorest performing older adults being unable to complete the study. However, older adults who did not complete the task were somewhat slower, but otherwise similar to the older adults who did complete the task. (The eight older adults who did not complete the study were of similar age (M = 70.1), $t(36) = 0.58$, $p = .568$ and had similar vocabulary scores (M = 22.4), $t(36) = 1.16$, $p = .254$, and years of education (M = 15.6), $t(36) = 0.99$, $p = .331$, as the older adults who did complete the study (Table 2). They were marginally slower on pattern comparison (M = 23.8), $t(36) = 2.01$, $p = .052$. Of the older adults who did not complete the task, the slowest performed only 10 trials. As a group, the nonfinishers performed numerically ($M = -6.51$), but not significantly worse than their finishing older adult counterparts ($M = -0.89$), $t(36) = 0.73$, $p = .468$, over the first 10 trials.)

Initial task representations were predictive of chemistry task performance for both age groups, supporting models of strategic choice (Bromme et al., 2009; Lovett & Schunn, 1999; Muis, 2007; Winne & Hadwin, 1998). However, young adults with accurate task representations generally outperformed older adults with similar task representations. Thus the evidence suggests a mental model utilization deficiency in older adults. This finding fits with age-related impairments in metacognitive control (Dunlosky et al., 2006; Dunlosky & Connor, 1997; Dunlosky & Hertzog, 2000; Hertzog, Dunlosky, et al., 2010; Hertzog et al., 2002, Hertzog, Sinclair, & Dunlosky, 2010). However, in our study these impairments were relatively small.

Although both age groups showed modest improvement in performance throughout the chemistry task, task representation updating (as measured by changes in C-scores) was rare in both age groups. This is not entirely surprising given findings that even younger adults do not fully update their strategy knowledge following task experience (Dunlosky & Hertzog, 2000; Hertzog et al., 2009; 2008; Hertzog, Price, & Dunlosky, 2012). Knowledge updating may have been additionally impaired if participants failed to understand or utilize the feedback provided during the chemistry task (Cassidy & Gutchess, 2015). It is also possible that improvements in task performance resulted primarily from learning *how much* of a chemical to add rather than learning new relationships between the chemical additives and the properties they affect. Pathfinder ratings capture only the presence or absence of relationships and not their degree or direction. An alternative would be to use a rolling regression technique that regresses the specific levels of each cue (solution property value) onto the strategic choices of the participant (decision to add a particular chemical; see Lagnado, Newell, Kahan, & Shanks, 2006 for an explanation). This technique estimates how each participant weighs the different cues throughout the task. Unfortunately rolling regression is not possible with the current data set because the cues have nonlinear and interactive relationships with each other and with other decisions on each trial. Given the complexity of the relationships within our task, and older adults documented struggles with learning new directional relationships (Chasseigne et al., 2004; Mata et al., 2010), it is all the more impressive that older adults obtained similar performance as that of younger adults. However, this performance equivalence should not be confused with strategy equivalence. Performance on the chemistry task is reflective of the appropriateness of the strategic choices. However, it does not indicate how participants improved their performance. Younger and older adults might have differed in whether they improved their performance through learning new relationships, fine-tuning relationships learned during the instructions phase, or considering multiple cues together rather than separately (which would not necessarily require new learning).

Despite the general lack of updating, when subtle improvements in task representations did occur, they were related to better task performance, supporting models of strategic choice (Bromme et al., 2009; Lovett & Schunn, 1999; Muis, 2007; Winne & Hadwin, 1998). Notably, older adults benefitted somewhat less compared with younger adults when they improved their task representations. This apparent inability to fully utilize their updated task representations indicates an age-related decrement in cognitive control and is somewhat consistent with both the metacognitive (Dunlosky & Hertzog, 2000; Hertzog et al., 2009, 2008, 2012) and probabilistic learning literatures (Chasseigne et al., 2004; Mata et al., 2010) where older adults struggle to learn and utilize complex relationships via monitoring and feedback. However, in the current studies, older adults' overall performance was still generally similar, if not superior to, that of younger adults.

General Discussion

Task representations are thought to guide strategic choices on a variety of tasks (Bromme et al., 2009; Lovett & Schunn, 1999; Muis, 2007; Winne & Hadwin, 1998). Despite known age differences in strategic choices, no previous study to our knowledge has examined age differences in task representations. Study 1 demonstrated that pathfinder-derived concept maps (C-scores) can be used to measure younger and older adults' task representations. Study 2 found a strong relationship between C-scores and performance on a novel chemistry task. These findings are consistent with the hypothesis that individual differences in task representation accuracy produce individual differences in performance (Bromme et al., 2009; Lovett & Schunn, 1999; Muis, 2007; Winne & Hadwin, 1998).

Neither study found significant age-related decrements in task representation formation. Because mental model formation is cognitively demanding (Schnotz & Preuss, 1997), it is important to consider why older adults are able to form accurate mental representations in our studies and others (e.g., Castel, 2007; Radvansky et al., 2003, 1990; Stine-Morrow et al., 2004, 2002). The self-paced nature of our and many other studies (but see Castel, 2007) may allow older adults to compensate for age-related declines by taking more time to learn and make decisions. Alternatively, older adults may have benefitted from their prior knowledge and schemas by incorporating the new task information into them. This seems unlikely given the novelty of the chemistry task, however older adults' vocabulary scores (a rough proxy for crystalized knowledge) correlated with chemistry task performance, $r = .41$.

Age differences in performance were also minimal. If anything, older adults performed slightly better than younger adults. This finding is consistent with research showing that older adults perform better when information can be incorporated into a coherent mental representation (e.g., Castle, 2007; Radvansky et al., 2003, 1990; Stine-Morrow et al., 2004, 2002). To more directly test this hypothesis, future research should manipulate whether the task relationships are provided with or without a unifying context (i.e., the chemistry task narrative). Additionally, future studies should include other measures for identifying task misrepresentations, such as critical error analysis (Harada, Mori, & Taniue, 2010) or strategic step skipping (Haider & Frensch, 1996; Kieras & Bovair, 1983).

Acknowledgments

We thank Kenzie Davis, Chandy Haskins, Eric Kaplan, Ashkon Koucheki, Amanda McCoy, Brent Nara, Kristin Nicholson, and Fizah Zafar for assistance with data collection.

References

Brigham M. C., & Pressley M. (1988). Cognitive monitoring and strategy choice in younger and older adults. Psychology and Aging , 3, 249–57. doi:10.1037/0882-7974.3.3.249

Bromme R. Pieschl S., & Stahl E. (2009). Epistemological beliefs are standards for adaptive learning: A functional theory about epistemological beliefs and metacognition. Metacognition and Learning , 5, 7–26. doi:10.1007/s11409-009-9053-5

Cassidy B. S., & Gutchess A. H. (2015). Age effects in adaptive criterion learning. The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences , 71, 1–8. doi:10.1093/geronb/gbv039

Castel A. D. (2007). Aging and memory for numerical information: The role of specificity and expertise in associative memory. The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences , 62, P194–P196. doi:10.1093/geronb/62.3.P194

Chasseigne G. Ligneau C. Grau S. Le Gall A. Roque M., & Mullet E. (2004). Aging and probabilistic learning in single- and multiple-cue tasks. Experimental Aging Research , 30, 23–45. doi:10.1080/03610730490251469

Curtis M. B., & Davis M. A. (2003). Assessing knowledge structure in accounting education: An application of Pathfinder Associative Networks. Journal of Accounting Education , 21, 185–195. doi:10.1016/S0748-5751(03)00024-1

Dunlosky J. Baker J. M. C. Rawson K. A., & Hertzog C. (2006). Does aging influence people's metacomprehension? Effects of processing ease on judgments of text learning. Psychology and Aging , 21, 390–400. doi:10.1037/0882-7974.21.2.390

Dunlosky J., & Connor L. T. (1997). Age differences in the allocation of study time account for age differences in memory performance. Memory and Cognition , 25, 691–700. doi:10.3758/BF03211311

Dunlosky J., & Hertzog C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. Psychology and Aging , 15, 462–474. doi:10.1037//0882-7974.15J.4

Gilbert D. K. Rogers W. A., & Samuelson M. E. (2004). Long-term retention of a spatial mental model for younger and older adults. Experimental Aging Research , 30, 217–224. doi:10.1080/03610730490274266

Goldsmith T. E. Johnson P. J., & Acton W. H. (1991). Assessing structural knowledge. Journal of Educational Psychology , 83, 88–96. doi:10.1037//0022-0663.83.1.88

Gomez R. L. Hadfield O. D., & Housner L. D. (1996). Conceptual maps and simulated teaching episodes as indicators of competence in teaching elementary mathematics. Journal of Educational Psychology , 88, 572–585. doi:10.1037//0022-0663.88.3.572

Haider H., & Frensch P. A. (1996). The role of information reduction in skill acquisition. Cognitive Psychology , 30, 304–37. doi:10.1006/cogp.1996.0009

Harada E. T. Mori K., & Taniue N. (2010). Cognitive aging and the usability of IT-based equipment: Learning is the key. Japanese Psychological Research , 52, 227–243. doi:10.1111/j.1468-5884.2010.00440.x

Hertzog C. Dunlosky J., & Sinclair S. M. (2010). Episodic feeling-of-knowing resolution derives from the quality of original encoding. Memory & Cognition , 38, 771–84. doi:10.3758/MC.38.6.771

Hertzog C. Kidder D. P. Powell-Moman A., & Dunlosky J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? Psychology and Aging , 17, 209–225. doi:10.1037//0882-7974.17.2.209

Hertzog C Price J Burpee A Frentzel W. J Feldstein S, & Dunlosky J. (2009). Why do people show minimal knowledge updating with task experience: Inferential deficit or experimental artifact? Quarterly Journal of Experimental Psychology , 62, 155–173. doi:10.1080/17470210701855520

Hertzog C. Price J., & Dunlosky J. (2008). How is knowledge generated about memory encoding strategy effectiveness? Learning and Individual Differences , 18, 430–445. doi:10.1016/j.lindif.2007.12.002

Hertzog C. Price J., & Dunlosky J. (2012). Age differences in the effects of experimenter-instructed versus self-generated strategy use. Experimental Aging Research , 38, 42–62. doi:10.1080/0361073X.2012.637005

Hertzog C. Sinclair S. M., & Dunlosky J. (2010). Age differences in the monitoring of learning: Cross-sectional evidence of spared resolution across the adult life span. Developmental Psychology , 46, 939–948. doi:10.1037/a0019812

Hines J. C. Hertzog C., & Touron D. R. (2015). Younger and older adults weigh multiple cues in a similar manner to generate judgments of learning. Aging, Neuropsychology, and Cognition , 22, 1–19. doi:10.1080/13825585.2015.1028884

Interlink. (1990). Rate . Las Cruces, NM: Interlink Inc.

Johnson P. J. Goldsmith T. E., & Teague K. W. (1994). Locus of the predictive advantage in Pathfinder-based representations of classroom knowledge. Journal of Educational Psychology , 86, 617–626. doi:10.1037//0022-0663.86.4.617

Kieras D. E., & Bovair S. (1983). The role of a mental model in learning to operate a device. Cognitive Science , 8, 255–273. doi:10.1207/s15516709cog0803_3

Kuhlmann B. G., & Touron D. R. (2011). Older adults' use of metacognitive knowledge in source monitoring: Spared monitoring but impaired control. Psychology and Aging , 26, 143–149. doi:10.1037/a0021055

Lagnado D. a. Newell B. R. Kahan S., & Shanks D. R. (2006). Insight and strategy in multiple-cue learning. Journal of Experimental Psychology: General , 135, 162–183. doi:10.1037/0096-3445.135.2.162

Lemaire P. (2010). Cognitive strategy variations during aging. Current Directions in Psychological Science , 19, 363–369. doi:10.1177/0963721410390354

Littell R. C. Milliken G. A. Stroup W. W., & Wolfinger R. D. (2000). SAS system for mixed models (4th ed.). Cary, NC: SAS Institute.

Lovett M. C., & Schunn C. D. (1999). Task representations, strategy variability, and base-rate neglect. Journal of Experimental Psychology: General , 128, 107–130. doi:10.1037//0096-3445.128.2.107

Mata R. von Helversen B., & Rieskamp J. (2010). Learning to choose: Cognitive aging and strategy selection learning in decision making. Psychology and Aging , 25, 299–309. doi:10.1037/a0018923

Matvey G. Dunlosky J. Shaw R. J. Parks C., & Hertzog C. (2002). Age-related equivalence and deficit in knowledge updating of cue effectiveness. Psychology and Aging , 17, 589–597. doi:10.1037//0882-7974.17.4.589

Morrow D. G. Leirer V. Altieri P., & Fitzsimmons C. (1994a). Age differences in creating spatial models from narratives. Language and Cognitive Processes , 9, 203–220. doi:10.1080/01690969408402116

Morrow D. G. Leirer V. Altieri P., & Fitzsimmons C. (1994b). When expertise reduces age differences in performance. Psychology and Aging , 9, 134–148. doi:10.1037/0882-7974.9.1.134

Muis K. R. (2007). The role of epistemic beliefs in self-regulated learning. Educational Psychologist , 42, 173–190. doi:10.1080/00461520701416306

Mutter S. A., & Asriel M. W. (2016). Gist and generalization in young and older adults' causal learning. The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences , 0, 1–9. doi:10.1093/geronb/gbw026

Nelson T. O. & Narens L. (1990). Metamemory: A theoretical framework and new findings. The Psychology of Learning and Motivation: Advances in Research and Theory, 26, 125–173.

Novak J. D. (1990). Concept mapping: A useful tool for science education. Journal of Research in Science Teaching , 27, 937–949. doi:10.1002/tea.3660271003

Novak J. D., & Cañas A. J. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. Information Visualization , 5, 175–184. doi:10.1057/palgrave.ivs.9500126

Price J. Hertzog C., & Dunlosky J. (2008). Age-related differences in strategy knowledge updating: Blocked testing produces greater improvements in metacognitive accuracy for younger than older adults. Aging, Neuropsychology, and Cognition , 15, 601–626. doi:10.1080/13825580801956225

Radvansky G. A. (1999a). Aging, memory, and comprehension. Current Directions in Psychological Science , 8, 49–53. doi:10.1111/1467-8721.00012

Radvansky G. A. (1999b). Memory retrieval and suppression: The inhibition of situation models. Journal of Experimental Psychology: General , 128, 563–579. doi:10.1037/0096-3445.128.4.563

Radvansky G. A. Copeland D. E. Berish D. E., & Dijkstra K. (2003). Aging and situation model updating. Aging, Neuropsychology, and Cognition , 10, 158–166. doi:10.1076/anec.10.2.158.14459

Radvansky G. A., & Dijkstra K. (2007). Aging and situation model processing. Psychonomic Bulletin & Review , 14, 1027–42. doi:10.3758/BF03193088

Radvansky G. A. Gerard L. D. Zacks R. T., & Hasher L. (1990). Younger and older adults' use of mental models as representations for text materials. Psychology and Aging , 5, 209–214. doi:10.1037/0882-7974.5.2.209

Radvansky G. A. Zacks R. T., & Hasher L. (1996). Fact retrieval in younger and older adults: The role of mental models. Psychology and Aging , 11, 258–71. doi:10.1037/0882-7974.11.2.258

Rowe A. L., & Cooke N. J. (1995). Measuring mental models: Choosing the right tools for the job. Human Resource Development Quarterly , 6, 243–255. doi:10.1002/hrdq.3920060303

Schneider W. Eschman A., & Zuccolotto A. (2002). E-Prime user's guide . Pittsburgh, PA: Psychology Software Tools, Inc.

Schnotz W., & Preuss A. (1997). Task-dependent construction of mental models as a basis for conceptual change. European Journal of Psychology of Education , 12, 185–211.

Staggers N., & Norico A. F. (1993). Mental models: Concepts for human-computer interaction research. International Journal of Man-Machine Studies , 38, 587–605. doi:10.17226/790

Stine-Morrow E. A. L. Gagne D. D. Morrow D. G., & DeWall B. H. (2004). Age differences in rereading. Memory & Cognition , 32, 696–710. doi:10.1177/0165025407084050

Stine-Morrow E. A. L. Morrow D. G., & Leno R. (2002). Aging and the representation of spatial situations in narrative understanding. The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences , 57, 291–297. doi:10.1093/geronb/57.4

Touron D. R., & Hertzog C. (2004a). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. Psychology and Aging , 19, 452–466. doi:10.1037/0882-7974.19.3.452

Touron D. R., & Hertzog C. (2004b). Strategy shift affordance and strategy choice in young and older adults. Memory & Cognition , 32, 298–310. doi:10.3758/BF03196860

Van Boven L., & Thompson L. (2003). A look into the mind of the negotiator: Mental models in negotiation. Group Processes & Intergroup Relations , 6, 387–404. doi:10.1177/13684302030064005

Winne P. H., & Hadwin A. (1998). Studying as self-regulated learning. In J. Dunlosky D. J. Hacker, & A. C. Graesser (Eds.), Metacognition in educational theory and practice (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum Associates.