Contents lists available at ScienceDirect



Journal of Memory and Language





Adults with poor reading skills, older adults, and college students: The meanings they understand during reading using a diffusion model analysis



Gail McKoon*, Roger Ratcliff

The Ohio State University, United States

ARTICLE INFO	A B S T R A C T
Keywords:	When a word is read in a text, the aspects of its meanings that are encoded should be those relevant to the text
Diffusion model	and not those that are irrelevant. We tested whether older adults, college students, and adults with poor literacy
Response time and accuracy	skills accomplish contextually relevant encoding. Participants read short stories, which were followed by true/
Context effects in meaning Older adults	false test sentences. Among these were sentences that matched the relevant meaning of a word in a story and sentences that matched a different meaning. We measured the speed and accuracy of responses to the test
Low literacy adults	sentences and used a decision model to separate the information that a reader encodes from the reader's speed/

lead to further understanding of reading skills.

Introduction

A complete understanding of a text requires understanding the meanings of its words as they are relevant to it. The full meaning of a word is not necessary and irrelevant aspects of its meaning are not useful and may be detrimental. Contextually relevant meanings are needed for ambiguous words such as "bug," but also for words that are not ambiguous. For a text that mentions a tomato, the relevant aspects of its meaning are different when the text discusses its color than when it discusses its shape. Tomatoes are red is relevant in the context of painting a picture of a tomato but tomatoes are round is relevant in the context of rolling a tomato across a table.

Research on reading is often organized by Perfetti and Stafura's (2014; Perfetti, 2007; also Cromley & Azevedo, 2007) general framework for comprehension. Word knowledge is at the center of this framework. This is reflected in the "lexical quality hypothesis" that skilled reading depends on high-quality, robust knowledge of words (Perfetti & Hart, 2002). High-quality representations are said to be essential to understanding which aspects of a word's meaning are contextually relevant and therefore essential to successful comprehension (Perfetti & Hart, 2002; Perfetti, 2007).

We examined the representations of words' meanings as they were encoded into short stories by testing memory for them. There was a series of blocks of trials and for each block, participants read six stories, one at a time. They then received a list of true/false test sentences, some for which the truth or falsity could be evaluated only by reference to one of the six stories and some for which the truth or falsity could be evaluated by reference to general knowledge (e.g., "the sky is concrete," "newspapers are reading material," "tomatoes are red," "tomatoes are round"). The materials of interest were pairs of stories like these "tomatoes" stories:

accuracy tradeoff settings. We found that all three groups encoded meanings as contextually relevant. The findings illustrate how a decision-making model combined with tests of particular comprehension processes can

- "This painting would require great accuracy. The painter searched many stores to find the color most suited to use in painting the ripe tomato."
- "The child psychologist watched the infant play. The little girl found a tomato to roll across the floor with her nose."

For the first, the meaning of tomato is more to do with red than round and for the second, it is the reverse. The test sentence "tomatoes are red" matches the first text and not the second, and the test sentence "tomatoes are round" matches the second but not the first. The question was whether verification of a matching sentence was easier than verification of a mismatching sentence.

The diffusion decision model

An important difference between our study and many other studies of comprehension including those that have investigated contextually relevant meaning is that we use a quantitative model of decision making (Ratcliff, 1978; Ratcliff & McKoon, 2008). For the experiments reported here, participants were asked to make two-choice decisions

https://doi.org/10.1016/j.jml.2018.05.005 Received 14 March 2017; Received in revised form 4 May 2018 Available online 07 June 2018 0749-596X/ © 2018 Published by Elsevier Inc.

^{*} Corresponding author at: Department of Psychology, The Ohio State University, Columbus, OH 43210, United States. *E-mail addresses*: mckoon.1@osu.edu (G. McKoon), ratcliff.22@osu.edu (R. Ratcliff).

(true-false).

The model solves problems that have been ubiquitous in decision making research. First, it translates accuracy and RTs into the same underlying components of the decision process, one of which is the information from memory upon which the true-false decisions are based. This allows direct measurement of the degree to which matching versus mismatching information is encoded into the meanings of the texts. Second, it explicitly addresses the speed/accuracy tradeoffs that individuals adopt, which is required because an individual may respond with low accuracy to test items because he or she does not know the relevant information or because he or she does know the information but decides to prioritize speed over accuracy. The criterion that an individual sets to determine the accuracy and RTs of his or her responses is a second component of the model, independent of the first. The third problem that the model addresses is that of scaling. For example, suppose the baseline mean RT for one group of individuals is 1000 ms and for another it is 700 ms. Further, suppose the mean difference in RTs between two experimental conditions is 150 ms for the first group and 80 ms for the other. The problem is how to interpret the effect of the independent variable. In other words, is the size of the effect the same for 150 out of 1000 as for 80 out of 700? The model allows this question to be answered. The fourth advantage of the model is that it can substantially increase the power to observe differences among the conditions of an experiment when there are relatively small numbers of observations and/or the variability in accuracy and RTs is too large to detect effects.

These contributions of the diffusion model to research on text comprehension are crucial for the development of theories about comprehension in general and about the comprehension skills of older and low-literacy individuals in particular. To develop such a theory, it is essential to know to what extent individuals have encoded textual information but this is not possible without separating away individuals' choices of speed/accuracy settings. It is also impossible to compare one individual's skills, or one population's skills, to another's without dealing with scaling issues. And, increasing the power of experiments should allow finer discriminations among theories.

In earlier research, the insights gained by application of the diffusion model have been demonstrated for elderly adults in a number of tasks, including lexical decision, recognition memory, numerosity judgments, and perceptual tasks. Generally, prior to application of the diffusion model, it had been assumed that all, or almost all, cognitive processes slow with age and so less information can be brought to bear on decisions. However, application of the model showed this is not the case for many tasks (e.g., Ratcliff, Thapar, & McKoon, 2004, 2010, 2011): elderly adults are slower than young adults in large part because they set their speed/accuracy settings to value accuracy more than young adults do.

The three experiments

We tested whether words' meanings were encoded in a contextually relevant fashion in three experiments. For Experiment 1, the participants were college students, for Experiment 2, they were older adults with a mean age of 70.6, and for Experiment 3, they were adults who read at only about the seventh grade level. McKoon and Ratcliff (1988) found that college students show contextually relevant encoding in an experiment similar to that reported here and so we expected to replicate their finding. The more important questions were whether this ability declines with age and whether it is an ability missing from the comprehension skills of poor readers. If word knowledge is indeed at the center of comprehension, then if older adults fail to properly encode it, there are serious implications for their abilities to use, for example, financial or health information. If poor readers fail, constructing meanings in context should be a focus of classes aimed at improving their reading skills.

Older adults

As we describe in more detail later, studies of context effects with older adults have produced mixed results, some indicating that they do encode meanings in a contextually relevant fashion and some that they do not. There have been many hypotheses about the effects of age on cognitive processes that are relevant to reading comprehension. Older adults may have a deficit in short-term memory, in binding one piece of information to another, in processing speed, and/or in the ability to discard irrelevant information. Especially relevant here, Craik and Byrd (1982) proposed that older adults are limited in the richness with which they encode words' meanings and therefore limited in the extent to which context determines what they encode.

In Experiment 2, we found that the older adults did encode contextually relevant meanings. Their responses were slower than those of the young adults but, as with the earlier research mentioned above, this was due, in the main, to differences in their speed-accuracy criteria settings. Older adults generally set their boundaries to strongly value accuracy, setting them within only a few percent of the boundaries that would give the maximum possible accuracy (Starns & Ratcliff, 2010; Starns & Ratcliff, 2012).

Adults with poor reading skills

Low literacy is a dramatically large problem in the United States (The National Center for Education Statistics; Baer, Kutner, & Sabatini, 2009; Kutner, Greenberg, & Baer, 2006; Miller, McCardle, & Hernandez, 2010; Greenberg, 2008). The International Adult Literacy Survey Institute. (2011) found that about 23% of adults in the United States read prose at the lowest level scored, indicating difficulty with comprehending even the most basic textual information; the National Assessment of Adult Literacy (Kutner et al., 2006) found that 43% lack the necessary literacy skills for most living wage jobs; and the Organization for Economic Co-operation and Development (OECD, 2013) found that one in six adults, about 36 million (two-thirds of them born in the United States) have low literacy skills. As Nicholas Kristof of the New York Times has put it (October 26, 2014), these data "should be a shock" to all Americans.

Previous research with poor readers has been interpreted as showing that they have difficulties with most, if not all, elements of comprehension, including establishing contextually relevant meanings. However, as with the older adults, we found that they did encode such meanings. And, again, their responses were slower than those of the young adults mainly because they set their boundaries farther apart, valuing accuracy more than speed.

College students: Experiment 1

Many studies have been interpreted as showing that good readers' understandings of word meanings during reading are context dependent. For college-age adults who read well, Barsalou (1982), Tabossi and Johnson-Laird (1980), Tabossi (1982), and McKoon and Ratcliff (1988) all found that the time taken to verify a true sentence was shorter when the information in the sentence was relevant to the information in an immediately preceding text. For example, in Tabossi's experiment, the time to verify that banks contain money was shorter after "the bank was robbed by three bandits" than after "the bank was built 10 years ago."

For good readers, context-dependent meanings appear not only to be available immediately after a context sentence but also encoded into memory with their context sentences. Anderson et al. (1976) and Anderson and Ortony (1975) used cued recall paradigms. Anderson et al. had participants study sentences (in mimeographed booklets) like "the fish attacked the swimmer" and "the fish avoided the swimmer" and found that "shark" was a better cue for the "attack" sentence than for the "avoid" sentence and Anderson and Ortony (1975) found that "basket" was a better cue for "the container held the apples" than "the container held the coca cola." McKoon and Ratcliff (1988) asked participants to decide whether test sentences were true or false and, after several stories had been read, found facilitation in RTs and accuracy when a test sentence was relevant to a text compared to when it was not.

Perhaps the most compelling finding about context-dependent meaning for good readers was provided by Roth and Shoben (1983) who showed that context can restructure the representation of a word's meaning. They used context sentences like "during the midmorning break, the two secretaries gossiped as they drank the beverage" and "before starting his day, the truck driver had the beverage and a donut at the truck stop." For both contexts, participants read an immediately following sentence faster if it instantiated "coffee" as the beverage, compared to a control sentence, implying that "coffee" was contextually relevant for both. However, if the context sentence was "during the midmorning break, the two secretaries gossiped as they drank the beverage," a sentence instantiating "beverage" as "tea" was read faster than a sentence instantiating it as "milk." The reverse pattern was obtained when the context sentence was "before starting his day, the truck driver had the beverage and a donut at the truck stop;" an instantiation of "milk" was read faster than an instantiation of "tea." These results indicate that "beverage" is understood in one case as coffee-or-tea-type beverages and in the other as coffee-or-milk-type beverages.

From results like these, we expected that college students would show contextually relevant encoding with the "tomatoes" materials.

Method

Participants

There were 32 participants, all receiving credit for an introductory psychology class at Ohio State University.

Materials

The materials were similar to those used by McKoon and Ratcliff (1988). There were 80 pairs of stories like the tomatoes example, one making one aspect of the meaning of a noun more relevant and one making another more relevant, each pair with two test sentences, making four conditions, as exemplified by the tomatoes stories: "Tomatoes are red" and "Tomatoes are round" were tested after "The child psychologist watched the infant play. The little girl found a tomato to roll across the floor with her nose." and after "This painting would require great accuracy. The painter searched many stores to find the color most suited to use in painting the ripe tomato." Other examples are "turtles have shells," "turtles are slow;" "a lamp is a piece of furniture," "a lamp gives light;" "airplanes have pilots;" and "airplanes carry passengers". For each of the two stories of a pair, there were also one sentence that was true according to the story (e.g., "the painter searched many stores") and one that was false (e.g., "the painting

Table 1

Response probabilities and median RTs.

required little accuracy"). In addition, there were 40 filler stories, each with one test sentence true according to the story and one false. There were also 80 test sentences that were true and 120 that were false according to general knowledge, all unrelated to any of the stories in the experiment. Examples are "cradles hold teenagers," "ants are mammals," "ribbons are delicious," "clocks are mysterious," "Mars is a planet," "beavers build dams," "cows moo," and "contact lenses correct vision."

Procedure

The stories and test sentences were presented on a PC screen and responses were recorded from its keyboard. To familiarize participants with the screen and keyboard, the experiment began with 32 items tested in lexical decision (half words and half nonwords).

Following the lexical decision items, there were 25 blocks of stories and test sentences, the first one for practice. To begin each block, participants pressed the space bar on the keyboard. Each block after the practice was made up of six stories, four experimental and two filler, in random order, and 26 test sentences. For stories with two lines, presentation time was 6 s, for three-line stories it was 9 s, and for four-line stories it was 11 s.

Each test list was made up of the target sentence for each experimental story, the true and false test sentences for each experimental story, the true and false test sentences for the filler stories, six sentences false by general knowledge, and four sentences true by general knowledge. The test sentences were in random order except that the target sentence for an experimental story was immediately preceded by the true test sentence for that story.

Each test sentence was displayed until a participant made a response, pressing the ?/ key for "true" and the zZ key for "false." If the response was correct, the screen was cleared for 50 ms and then the next test sentence was displayed. If the response was not correct, the message "ERROR" was displayed for 900 ms, then the screen was cleared for 50 ms and the next test sentence displayed. Participants were asked to respond as quickly and accurately as possible.

Design

There were the four conditions described above: the first story of a pair was presented and the target test sentence matched that story or it mismatched that story (i.e., it matched the other story of the pair), or the second story of a pair was presented and the target sentence matched or mismatched it. For the 32 participants, a Latin-square counterbalanced the four conditions giving 8 participants for each condition and 20 pairs of stories for each condition.

Results: RTs and accuracy

For all three experiments, responses longer than 4000 ms and shorter than 350 ms (about 1% of the data) were eliminated from

	ABLE students			College stu	idents		Older adults		
Condition and (number of observations per participant)	Pr "True"	Median RT "True"	Median RT "False"	Pr "True"	Median RT "True"	Median RT "False"	Pr "True"	Median RT "True"	Median RT "False"
Matching (40)	0.951	1575	1796	0.982	1088	1346	0.976	1396	1941
Mismatching (40)	0.930	1686	1944	0.957	1167	1358	0.940	1482	1946
True from Matching-Mismatching Story (80)	0.827	2184	2678	0.904	1538	1841	0.925	1948	2370
False from Matching-Mismatching Story (80)	0.539	2389	2695	0.281	1778	1733	0.267	2192	2244
True from Filler Story (40)	0.787	2046	2370	0.904	1431	1782	0.913	1861	2351
False from Filler Story (40)	0.516	2313	2570	0.288	1716	1701	0.293	2156	2247
True from General Knowledge (80)	0.820	1937	2133	0.940	1378	1513	0.920	1697	1915
False from General Knowledge (120)	0.150	2413	2273	0.052	1564	1495	0.051	2254	1840

analyses. Table 1 shows median RTs and accuracy averaged over participants. (We use median RTs because medians are used in fitting the diffusion model to the data – the 0.5 quantile RTs.) There were no significant differences between the two stories of the pairs and so the two matching conditions were combined and the two mismatching conditions were combined for analyses of the data. This was true for all three experiments. We present results from standard analyses and then mixed effects t-tests and ANOVAs.

RTs for matching test sentences were shorter than RTs for mismatching ones (t(31) = -5.3, p < .05, mixed effect model t = -7.3, p < .05) with a mean difference of 67 ms (95% CI 42–94 ms). They were also more accurate (t(31) = 3.2, p < .05, mixed effect model t = 4.1, p < .05), with a mean difference of 0.025 (95% CI 0.009–0.041). This is a robust result that generalizes McKoon and Ratcliff's (1988) finding to new stories and more of them (48 vs. 80).

Older adults: Experiment 2

To understand how age affects comprehension processes, it is necessary to understand the extent to which older individuals encode textual information: this requires separating that information from speed/accuracy settings. To our knowledge, there have been no studies that have used diffusion models to address this question for these kinds of materials. A finding of significant decline would have important consequences for theories about the effects of age on comprehension. It would also have serious, practical, implications for how textual information should be constructed for older adults such that they can achieve full comprehension. In addition, a finding of no significant decline would provide a baseline against which individuals with Mild Cognitive Impairment (MCI) and early Alzheimer's Disease (AD) could be compared. It might be that appropriately establishing the meanings of words is one of the last comprehension skills to be lost, which could suggest possibilities for tailoring verbal interactions with MCI and AD individuals.

The volume of research investigating older adults' word knowledge is extremely large and many specific issues have been addressed. All of it, as we review, indicates that older adults have the knowledge needed to understand which of the meanings of a word are relevant to a text being read. We first review this literature and then discuss whether older adults actually do instantiate meaning as contextually relevant.

Crystallized intelligence is defined as semantic knowledge that is preserved with age and it includes semantic information about words. The preservation of various types of crystallized knowledge, including word knowledge, across age has been demonstrated in many studies (e.g., Cattell, 1963; Labouvie-Vief, 1977; Schaie, 1970; Schretlen et al., 2000; Brod, Werkle-Bergner, & Shing, 2013; Horn & McArdle, 1992). It contrasts with fluid intelligence, which is the ability to process, make use of, and reason with new information, and fluid intelligence does decline with age.

Vocabulary – the number of words an individual knows – is one measure of word knowledge. Many studies have shown that it does not decrease with age and might even increase, perhaps reflecting the much longer experience with words that older adults have compared to young adults (e.g., Ackerman & Rolfhus, 1999; Bowles, Grimm, & McArdle, 2005; Bowles & Salthouse, 2008; Burke & Shafto, 2008; Lezak, Howieson, Loring, Hannay, & Fischer, 2004; Kave & Halamish, 2015; Kave & Yafe, 2014; Kemper & Sumner, 2001; Ronnlund, Nyberg, Backman, & Nilsson, 2005; Salthouse, 1993; Uttl, 2002; Verhaeghen, 2003).

Beyond the number of words older adults know, it appears that there is no decrement in knowledge of their meanings. Studies have shown that semantic knowledge is well preserved (e.g., Ackerman & Rolfhus, 1999; Allen, Sliwinski, & Bowie, 2002; Goral, Spiro, Albert, Obler, & Tabor Connor, 2007; Light & Albertson, 1988; Salthouse, 2009) and that there is no decrement in the organization of the meanings of words, including associations among words, associations of a word to its synonyms, and associations of a word to its antonyms (e.g., Burke & Peters, 1986; Lovelace & Cooley, 1982; Scialfa & Margolis, 1986). Older adults also show semantic priming effects (e.g., in lexical decision, Balota & Duchek, 1988; Burke & Harrold, 1988; Burke, White, & Diaz, 1987; Chiarello, Church, & Hoyer, 1985; Duchek & Balota, 1993; Howard, McAndrews, & Lasaga, 1981; Laver, 2009; Laver & Burke, 1993; Myerson, Ferraro, Hale, & Lima, 1992; Radvansky, Gibson, & McNerney, 2014; White & Abrams, 2004) and they are able to improve their recall of word pairs if there are semantic associations between the words (Badham, Estes, & Maylor, 2012; Naveh-Benjamin, Craik, Guez, & Kreuger, 2005; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003).

To encode meanings as they are relevant to a text requires not only knowledge of individual words and their meanings but also general knowledge of the world. For example, knowing what painters do and why people go shopping is needed for understanding the tomato stories. Much knowledge of the world is crystallized knowledge and so preservation with age would be expected. Many studies have shown this (e.g., Arbuckle, Vanderleck, Harsany, & Lapidus, 1990; Brod et al., 2013; Cattell, 1963; Cornelius & Caspi, 1987; Dixon, 2003; Hess, 1985; Labouvie-Vief, 1977; Park & Schwarz, 2000; Radvansky & Dijkstra, 2007; Salthouse, 2014; Schaie, 1970; Schretlen et al., 2000; Staudinger, Cornelius, & Baltes, 1989). Particularly relevant to comprehending words as they are contextually relevant are schemas (McKoon, Ratcliff, & Seifert, 1989; Schank & Abelson, 1977; Seifert, McKoon, Abelson, & Ratcliff, 1986). Schemas are representations in semantic memory of sets of typical characters, actions, and outcomes that occur in real-life episodes. "Eating out at a restaurant" would be a schema that includes, for example, information about waiters, chefs, silverware, plates, glasses, food, salt, drinks, menus, bills, money, and receipts, plus all the relations among these characters and objects and how the activities occur through time. In schemas, the meanings of words are dependent on the context in which they are used. For example, for "silverware," its meaning in the schema would depend on context, real silverware in an expensive restaurant, plastic in MacDonald's. Preserved schema knowledge for older adults has been demonstrated by Arbuckle et al. (1990), Radvansky and Dijkstra (2007), Miller, Stine-Morrow, Kirkorian, and Conroy (2004), and Stine-Morrow, Miller, and Hertzog (2006). Radvansky and Dijkstra (2007; p. 1036) concluded that "the range of knowledge that is available to people as they are actively processing information online is essentially the same in younger and older adults."

All of the data just reviewed indicate that older adults' knowledge of words and of the world is well-preserved. From this, the obvious prediction is that they can use this knowledge to incorporate pieces of textual information into a whole in which all the meanings are contextually relevant to each other. On the other hand, there are many difficulties for older adults that could plausibly prevent them from instantiating only contextually relevant meanings. These include short-term memory deficits (e.g., Craik & Byrd, 1982), limitations on attentional resources (e.g., Craik, 1983; Stine-Morrow et al., 2006), deficits in abilities to bind one piece of information to another (Naveh-Benjamin, Guez, Kilb, & Reedy, 2004), decreases in processing speed (e.g., Salthouse, 1996), and deficits in abilities to discard irrelevant information as they read (Hamm & Hasher, 1992).

As described above, the results of some studies of context effects with older adults have been interpreted as showing that they do encode meanings in a contextually relevant fashion and other studies that they do not. We have already described one positive finding, priming in lexical decision and naming tasks, and there are a number of other positive findings. Madden (1988; also Hopkins, Kellas, & Paul, 1995) had participants make lexical decisions about the final words of sentences with the final words being either congruous or incongruous in the context of their sentences. When the final words were visually degraded, RTs benefited from congruent context more for older than young adults (this could be a scaling effect, see the discussion in

McKoon and Ratcliff (2012, p 421)). Wingfield, Aberdeen, and Stine (1991) and Pichora-Fuller, Schneider, and Daneman (1995) found that older adults benefited more from supportive context when asked to name words heard in noise than did young adults. Wingfield, Alexander, and Cavigelli (1994) asked participants to pronounce single words heard in noise. If they failed, a word of context was added; if they failed again, another word was added, and so on. Older adults improved with additional words as much as young adults. Cohen and Faulkner (1983) asked participants to make a lexical decision about the final word of a sentence. When the word did not fit the context well, older participants were slower than young participants, but when it did, facilitation was larger for the older participants. Rayner, Reichle, Stroud, Williams, and Pollatsek (2006: also Kliegl, Grabner, Rolfs, & Engbert, 2004) measured fixation times on words that did or did not fit their context and found that the difference between the two contexts was not significantly different.

The studies just reviewed are countered by others that show deficits in older adults' contextual encodings. Craik and Byrd's (1982) hypothesis was that older adults' decreasing cognitive resources result in decrement of the extensiveness and depth of information they encode during reading, with the result that the information they encode is "less modified by the specific context in which it occurs" than is the case for young adults (p. 208; see also Light, Valencia-Laver, & Zavis, 1991). In accord with this, Dagerman, Macdonald, and Harm (2006) had participants listen to sentences up to an ambiguous word that occurred just before the last word. The last word disambiguated the ambiguous word and it was presented visually for naming. Young adults, but not older adults, were faster at naming the disambiguating word when it matched the context of the sentence than when it did not. In a large body of research (Federmeier, 2007; Federmeier & Kutas, 2005; Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, & Kutas, 2002; Wlotko, Federmeier, Fabiani, & Gratton, 2011; Wlotko & Federmeier, 2012; Wlotko, Federmeier, & Kutas, 2012; Wlotko, Lee, & Federmeier, 2010), Federmeier and colleagues have examined ERP signals, in particular the N400, as indicators of a reader's use of context information. Older adults made less use of context than young adults, as shown by later and smaller N400 signals. (These results come with a strong and important caveat: the words are presented at a rate of 500 ms per word, considerably slower than words are usually read or spoken, and it might be that this is more disconcerting for the older than the young adults, and therefore it is this that leads to the N400 differences, not sentential context.)

Method

The were 32 participants, community-living adults with a mean age of 70.6 with a standard deviation of 6.2. None had been diagnosed as having MCI or AD. They were paid \$20 for the 45-min experiment.

The materials, procedure, and design were the same as for Experiment 1 and we analyzed the data in the same ways. The RT and accuracy data are shown in Table 1.

Results

Like the college students, the older adults showed the matching effect. They were more accurate (t(31) = 3.0, p < .05), mixed effect model t = 5.6, p < .05) with a mean difference of 0.036 (95% CI 0.011–0.059) for matching than mismatching sentences and their RTs were shorter (t(31) = -6.5, p < .05), mixed effect model t = -4.8, p < .05) with a mean difference of 91 ms (95% CI 63–120 ms). For all the conditions in the experiment, the older adults were slower but had similar accuracy to the college students (Table 1).

Experiment 3: Adults with poor reading skills

Low-literacy adults have difficulties with many elements of

comprehension. Studies have suggested difficulties with appropriately integrating pieces of information in a text with each other, integrating pieces of information with general knowledge, establishing appropriate causal connections among pieces of information, identifying the main ideas of a text, and establishing the referents of pronouns, and it has been suggested that difficulties like these can occur even when the reader has the requisite general knowledge with which to accomplish these processes. Results like these have been found in many studies, with a variety of paradigms, including studies by Barnes, Ahmed, Barth, and Francis (2015), Bowyer-Crane and Snowling (2005), Cain and Oakhill (1999), Cain and Oakhill (2006), Cain, Oakhill, Barnes, and Brvant (2001), Cain, Oakhill, and Lemmon (2004), Garnham, Oakhill, and Johnson-Laird (1982), Laing and Kamhi (2003), Magliano and Millis (2003), Long and Golding (1993), Long, Oppy, and Seely (1994), Oakhill (1982), Oakhill (1983, Oakhill (1984, Oakhill (1993, Oakhill (1994), Oakhill and Yuill (1986), Oakhill, Yuill, and Donaldson (1990), Oakhill, Yuill, and Parkin (1988), Singer, Andrusiak, Reisdorf, and Black (1992), Todaro, Millis, and Dandotkar (2010); Whitney, Ritchie, and Clark (1991); and Yuill, Oakhill, and Parkin (1989).

The available literature also suggests that they may have difficulties with encoding meaning in a contextually relevant fashion. Gernsbacher, Varner, and Faust (1990) had participants read sentences with ambiguous final words, each followed by a test word for which participants were to decide whether or not it matched the meaning of the sentence. At 850 ms after the end of the sentence, skilled readers responded faster (relative to a control) to words representing appropriate meanings than words representing inappropriate meanings; poor readers did not. Henderson, Snowling, and Clarke (2013) had participants name pictures that were consistent or inconsistent with a preceding sentence. Skilled readers named consistent pictures faster (relative to a control) than inconsistent pictures but poor readers did not. Merrill, Sperber, and McCauley (1981) used a Stroop task; skilled readers showed interference only on target words that were relevant to the context of a preceding sentence whereas less-skilled readers showed interference on both relevant and irrelevant target words. Perfetti, Yang, and Schmalhofer (2008; also Yang, Perfetti, & Schmalhofer, 2005) found that low-literacy adults were less able to integrate words in a text with each other than skilled readers. Hannon and Daneman (2004) used Barton and Sanford's (1993) anomaly detection task in which participants read anomalous sentences like "The authorities were trying to decide where to bury the survivors" (of a plane crash); low-literacy readers were less likely to notice the anomaly.

Our intention for this experiment was to use the diffusion model approach in a manner analogous to the approach we used for older adults. Older adults' slow responses to stimuli were shown in many paradigms to come from wider boundaries, not poorer drift rates. Whether low-literacy adults' reading difficulties come from too narrow boundaries, slower nondecision processes, and/or poorer drift rates will significantly affect the development of literacy models.

Method

There were 32 participants who were paid \$20 for the 45-min experiment. They varied from 21.7 years old to 79.2 years old, with a mean of 44.9 and standard deviation of 13.7. Their reading levels were measured by the TABE test, a widely used test of adults' reading ability that is used in Adult Basic Literacy (ABLE) classes to determine reading grade levels. It consists of a series of texts, each with several paragraphs, on topics such as household cleaners, cell-phone purchasing plans, and "The Power of Color," with several multiple-choice questions for each. Scores on this test are translated into reading grade levels. For the ABLE students in our study, the students' grade levels varied from 3.8 to 10.7 with a mean of 6.9 and a standard deviation of 1.9. (Note that the TABE test seems quite reliable in the context here: McKoon & Ratcliff, 2016, showed that scores on the TABE were correlated with drift rates from a lexical decision task 0.48.)



Fig. 1. A: an illustration of the diffusion process. The panel shows three simulated paths with mean drift rate v, starting point z, and boundary separation a. One process hits the top boundary quickly, another hits it later, and another hits the bottom boundary in error. The panel also shows how the model predicts the right-skewed shapes of RT distributions: most processes hit the boundary quickly but some hit later. Drift rate is normally distributed across trials with SD η , starting point is uniformly distributed with range s_{z_2} and nondecision time is uniformly distributed with range s_r . B: an illustration of the mapping from RT distributions and accuracy to drift rates, boundary settings, and nondecision time. C: how an increase in boundary separation from a_1 to a_1 produces an increase in RT differences for a constant drift rate difference. D: how a constant difference in drift rates produces a larger RT difference if both drift rates are lower.

The materials, procedure, and design were the same as for the other experiments. We had constructed those materials with words familiar to the ABLE students, which we determined with a lexical decision task. All the content words were tested with ABLE students who did not participate in the experiment described here. For each of a series of strings of letters, they were asked to decide as quickly and accurately as possible whether the string was or was not an English word. Words for which accuracy was less than 80% were not used in the stories. While

lexical decision responses may not reflect the full meanings of words and they may not directly assess the possible "sluggishness" of wordlevel skills (e.g., Perfetti & Stafura, 2014; Perfetti & Stafura, 2015), they do require at least some level of familiarity with the words.

Results

Like the college students and the older adults, and surprisingly given the literature reviewed above, the ABLE students showed the matching effect. Their responses were more accurate for the matching sentences than the mismatching ones (t(31) = 2.5, p < .05), mixed effect model t = 2.1, p < .05) with a mean difference of 0.021 (95% CI 0.003–0.038) and they were faster (t(31) = -5.7, p < .05), mixed effect model t = -5.1, p < .05) with a mean difference of 110 ms (95% CI 71–150 ms).

Comparing results across the three groups of participants

We conducted analyses of variance to compare the three groups. The main effect of matching versus mismatching test sentences was significant in accuracy (F(1, 93) = 24.2, p < .05, $\eta^2 = 0.120$) and median RTs (F(1, 93) = 98.4, p < .05, $\eta^2 = .035$). Accuracy did not differ significantly across the three groups (F(2, 93) = 2.9, p > .05, $\eta^2 = .084$) but median RTs did (F(2, 93) = 44.5, p < .05, $\eta^2 = .943$). The differences between the two types of test sentences were not significantly different among the three groups for accuracy (F(2, 93) = 0.65, p > .05, $\eta^2 = .014$) or for median RTs (F(2, 93) = 1.83 (p > .05, $\eta^2 = .038$). With mixed effects modeling, for median RTs, the three groups differed (t = 9.3) and there was a matching effect (t = 10.3). For accuracy, there was an effect of group (t = -2.9) and there was a matching effect (t = -6.2).

The matching-mismatching differences cannot be directly interpreted because of scaling issues, especially for RTs. Combining the matching and mismatching conditions, median RT was 1128 ms for the college students, 1439 ms for the older adults, and 1692 ms for the ABLE students. The matching effects were 67 ms, 91 ms, and 110 ms for the three groups respectively. It is for reasons like this that a model is needed to integrate accuracy and RTs into one, underlying, measure of differences among experimental conditions.

The categories of test sentences other than the matching-mismatching ones (shown in Table 1) showed significant differences among the three groups of participants. For sentences true according to a story (we combined data for true sentences from experimental and filler stories weighted by their numbers of observations, 80 and 48): RTs, F(2, 93) = 31.77, p < .05, $\eta^2 = 0.405$, and accuracy, F(2, 93) = 19.11, p < .05, $\eta^2 = 0.291$; for sentences false according to a story (combined in the same way): RTs, F(2, 93) = 31.46, p < .05, $\eta^2 = 0.404$), and accuracy, F(2, 93) = 23.38, p < .05, $\eta^2 = 0.334$; for sentences true according to general knowledge: RTs, F(2, 93) = 32.81, p < .05, $\eta^2 = 0.413$, and accuracy, F(2, 93) = 16.83, p < .05, $\eta^2 = 0.265$; and for sentences false according to general knowledge: RTs, F(2, 93) = 12.78, p < .05, $\eta^2 = 0.215$ and accuracy, F(2, 93) = 39.63, p < .05, $\eta^2 = 0.460$.

The diffusion model

In the Introduction, we stressed that the speed/accuracy tradeoffs individuals select prevent direct measurement of the knowledge they bring to bear in making decisions. In the great majority of studies we reviewed for older adults and for adults with poor literacy skills, the dependent variables were accuracy or RT, but not both. There have been no applications of models to distinguish tradeoffs from knowledge, that is, extract decision related factors from knowledge.

The three experiments described in this article showed the utility of the diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008): it accounted for the data from the three experiments, at the level of individual participants, and it accomplished this for responses considerably slower than those that are typical for applications of the model (around 2.5 s vs 700 ms).

In this section, we describe the model. It is illustrated in Fig. 1. Total RT is the sum of the time taken by nondecision processes, which are those that encode a stimulus and those that execute a response, combined into the parameter T_{er} , and the time taken to make a decision. The top panel illustrates the decision process. The information from memory about a test sentence is accumulated over time from a starting point (z) to one or the other of two criterial amounts, or boundaries, one for "true" and one for "false." The stronger the information from memory, the higher the rate of accumulation, called drift rate, v. Drift rates for true sentences have positive values, moving from the starting point toward the upper boundary (a), and drift rates for false sentences have negative values, moving from the starting point toward the lower boundary (0). A response is executed when the amount of accumulated information reaches a boundary, 0 or a. The process of accumulating information is noisy. Three instances are shown in the figure. They have exactly the same mean drift rate but noise means that they approach the boundaries at different rates and that they sometimes reach the wrong boundary. This "within-trial" noise leads to the right-skewed distributions of RTs shown in the figure (Ratcliff, 1978; Ratcliff, 2002; Ratcliff & McKoon, 2008; Ratcliff, Van Zandt, & McKoon, 1999).

The stimuli of a particular condition in an experiment (e.g., contextually relevant test sentences, filler test sentences) are assumed to all have the same mean drift rate v. However, the value of drift rate for a particular condition is assumed to vary around its mean from trial to trial. This assumption comes from the notion that an individual cannot hold the mean value of drift rate exactly the same from one trial of a condition to the next (Ratcliff, 1978). For the same reason, it is assumed that there is trial-to-trial variability in the starting point (Laming, 1968) and in nondecision time (Ratcliff & Tuerlinckx, 2002).

The diffusion model has been validated many times, is applied in a wide number of domains (see Ratcliff & McKoon, 2008; Forstmann, Ratcliff, & Wagenmakers, 2016; Ratcliff, 2014; White, Ratcliff, Vasey, & McKoon, 2010), and explains all aspects of two-choice data: accuracy, mean RTs for correct responses, mean RTs for incorrect responses, the shapes and locations of RT distributions, and the relative speeds of correct and incorrect responses. As shown in Fig. 1, the model maps these data to underlying components of processing: drift rate (v), starting point (z), boundary separation (a), and nondecision time (T_{er}). In the figure, the data, on the left, are mapped through the model to give the values of the components, on the right.

The model's parameters are a drift rate for each condition in an experiment that is different in difficulty, the standard deviation in across-trial variability in drift rate, nondecision time, the range of nondecision time across trials, boundary separation, starting point, and the range of starting point across trials, plus a parameter to represent the proportion of responses that come from some process other than that of interest (e.g., lack of attention). We call these latter contaminant RTs and the proportion of them was less than 0.005 in the three experiments here.

We stress that the model is tightly constrained. The first and most powerful constraint comes from the requirement that the model fit the right-skewed shape of RT distributions that is almost always obtained (Ratcliff, 1978; Ratcliff, 2002; Ratcliff & McKoon, 2008; Ratcliff et al., 1999; Ratcliff, 2008). Second, across experimental conditions that vary in difficulty (and are randomly intermixed at test), changes in accuracy, RT distributions, and the relative speeds of correct and error responses must all be captured by changes in only one parameter of the model, drift rate. The boundaries cannot be adjusted as a function of difficulty because it would be necessary for the system to know which level of difficulty was being tested before the drift rate could be determined. Third, across experimental conditions that vary in speed/accuracy criteria (e.g., speed versus accuracy instructions), all the changes in accuracy, RT distributions, and the relative speeds of correct and error responses must be captured by changes in the settings of the response boundaries (and, empirically, there are also sometimes small changes in nondecision time).

Importantly, the model is identifiable and falsifiable. If data are simulated from the model and then the model is fit to the simulated data (accuracy and RTs), the parameters recovered from the data are within a percent or so of the parameters that were used for the simulation (e.g., Ratcliff & Childers, 2015; Ratcliff & Tuerlinckx, 2002). While it is relatively easy for the model to fit only mean RTs for correct responses and accuracy, and it can do so with a range of different parameter values (i.e., it would not be identifiable), it must also meet the three constraints just mentioned. Also, Ratcliff (2002) made up several sets of fake but quite plausible data and showed that the diffusion model failed (dramatically) to fit them. (Note that in most comparisons made so far, conclusions drawn from sequential sampling models other than the diffusion model, e.g., Usher & McClelland, 2001; Ratcliff, Thapar, Smith, & McKoon, 2005; Donkin, Brown, Heathcote, & Wagenmakers, 2011, have been the same as those from the diffusion model.)

The reason in the model that an individual's accuracy cannot be predicted from his or her speed or vice versa is that drift rates and boundary settings are separate from (and often independent of) each other. For example, for a given value of accuracy, an individual might set his or her boundaries close together and so respond quickly or he or she might set them farther apart and so respond slowly. It is this independence that allows the model to separate an individual's speed/ accuracy tradeoffs from the strength with which information has been encoded into memory. In the Discussion section below, we give examples of research that illustrates the power of the model to explain empirical data.

Fitting the model to the data

To fit the model to the data, the RT distributions were represented by 5 quantiles, the .1, .3, .5 (the median), .7, and .9 quantiles. The model was fit with a chi-square minimization method that is fully described in Ratcliff and Tuerlinckx (2002; Ratcliff & Childers, 2015). The correct and error RT distributions were weighted by the number of observations (because the chi-square method uses frequencies). The model was fit to each participant's data individually and it was fit to the eight categories of data shown in Table 1.

For all three experiments, there were 15 parameters for the model: a drift rate for each of the 8 conditions, boundary separation, starting point, nondecision time, the standard deviation in drift rate across trials, the ranges of variability in starting point and nondecision time across trials, and the proportion of contaminant RTs. The model parameters and SDs in the model parameters are shown in Table 2.

The first result is that the model fit the data well for all three experiments, as shown by the chi-square values in Table 2. The number of degrees of freedom was 73 (the number of conditions multiplied by 11 which is the number of bins for correct and error RTs between and outside the 5 quantiles minus 1 because the probabilities must add to 1 and minus the number of parameters). This gives a critical chi-square

value of 93.9 at the 0.05 level. The chi-square test is a very conservative test so, even when chi-square values are lower than twice the critical value (as a rule of thumb), the fit of the model to data is good (see Ratcliff, Thapar, Gomez et al., 2004, for a discussion of model fitting and chi-square values). For the older adults, college students, and ABLE students, only 1, 2, and 3 participants, respectively, had chi-square values greater than 2 times the critical value. The best-fitting values for all the model's parameters except drift rates are given in Table 2 and those for drift rates in Table 3.

The good fit of the model to the data for all three experiments is also shown in the plots in Figs. 2-4 where the x-axis is the data and the yaxis is the model's predictions for each experimental condition for each participant. Fig. 2 shows the plots for the ABLE students, Fig. 3 for the college students, and Fig. 4 for the older adults. There are about 250 data points for each plot, 32 participants with 8 conditions per participant. The x's represent conditions for which there were fewer than 25 and more than 6 observations (data points with less than 6 observations were not plotted) and the o's represent conditions for which there were more than 24 observations. The first plot in each figure shows the proportions of correct responses. For the college students and the older adults, there are only a few deviations greater than 5%. For the ABLE students, the model predicts lower values of accuracy than observed in the data for only a moderate number of responses. The other three plots in each figure show the 0.1, 0.5, and 0.9 quantiles of the RT distributions. For all of the participant groups, there are a modest number of cases for which the predicted 0.1 and 0.5 quantiles are lower than the data and for which the predicted 0.9 quantile is higher than the data. Many of the larger misses come from conditions with few responses (the x's in the plots).

For each of the three groups of participants, the matching minus mismatching difference in drift rates was significant. T-tests showed that for college students, t(31) = 5.0, p < .05, with a mean difference of 0.123 and 95% CI of 0.073–0.173; for older adults t(31) = 6.1, p < .05, with a mean difference of 0.159 and 95% CI of 0.106–0.213; and for ABLE students, t(31) = 3.9, p < .05, with a mean difference of 0.074 and 95% CI of 0.035-0.112.

Implicit in the good quality of the fit of the model to the data is that it separated those elements of responding that mainly determined RTs from the information that was available from memory. The differences in RTs between the three groups of participants were mainly due to differences in their boundary settings and nondecision times (Table 2). The differences among the groups were significant for both, F(2), 93) = 14.84, p < .05, η^2 = .242 for the distance between the two boundaries and F(2, 93) = 26.63, p < .05, η^2 = .364 for nondecision times.

With boundary settings and nondecision times abstracted away, the information encoded in memory, that is, drift rates, can be directly compared among the college students, the older adults, and the ABLE students (Table 3). We conducted a two-way analysis of variance with participant group and matching/mismatching as the factors. Both main effects were significant, F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = 14.48, p < .05, $\eta^2 = .355$, and F(2, 93) = .355, and F(2, 993) = 76.79, p < .05, η^2 = .218, respectively. The interaction was also significant, F(2, 93) = 3.36, p < .05, η^2 = .019, with the matching/

Table 2

M	leans	and	standar	d c	levia	tion	in	mod	lel	parameter	rs
---	-------	-----	---------	-----	-------	------	----	-----	-----	-----------	----

Acais and standard deviation in model parameters.											
Participant group		а	T _{er}	η	Sz	po	s _t	z	χ^2		
ABLE	Mean	0.263	1.260	0.088	0.060	0.003	0.788	0.150	135.8		
College students		0.213	0.905	0.091	0.039	0.004	0.575	0.124	107.0		
Older adults		0.242	1.200	0.097	0.033	0.004	0.670	0.143	106.0		
ABLE	SD	0.036	0.291	0.021	0.052	0.003	0.131	0.024			
College students		0.030	0.131	0.016	0.033	0.003	0.169	0.018			
Older adults		0.042	0.169	0.028	0.027	0.013	0.175	0.029			

Note. a = boundary separation, z = starting point, T_{er} = nondecision component of response time, η = standard deviation in drift across trials, s_z = range of the distribution of starting point (z), s_t = range of the distribution of nondecision times, p_o = the proportion of contaminants.

Table 3

Means and standard deviations in drift rates.

Participant group	Match	Mismatch	True from match- mismatch story	False from match- mismatch story	True from filler story	False from filler story	True from general knowledge	False from general knowledge
ABLE Mean	0.314	0.240	0.076	- 0.007	0.076	- 0.004	0.086	-0.116
College students	0.502	0.379	0.122	- 0.087	0.153	- 0.087	0.187	-0.223
Older adults	0.514	0.355	0.131	- 0.084	0.153	- 0.075	0.178	-0.236
ABLE SD	0.153	0.124	0.041	0.055	0.060	0.059	0.059	0.071
College students	0.150	0.171	0.037	0.054	0.060	0.057	0.062	0.073
Older adults	0.157	0.160	0.047	0.050	0.085	0.057	0.100	0.077



Fig. 2. Plots of accuracy and the .1, .5 (median), and .9 response time (RT) quantiles for data (x-axis) and predicted values from fits of the diffusion model (y-axis) for correct responses for ABLE students. The o's are for data with 25 or more observations and the x's are for between 7 and 24 observations. Note that the scales on this figure and Figs. 3 and 4 are different (in order to show the largest spread of points in each plot).

mismatching difference smaller for the ABLE students than the college students and older adults. This finding for the ABLE students could be interpreted as a deficit for them but it could also be a baseline effect; the percentage differences for matching-mismatching were 31%, 32%, and 45% for the ABLE students, college students, and older adults, respectively. If it is a baseline effect, then the ABLE students show no proportional deficit. For the current experiments, it is not possible to decide between these explanations.

In the presentation of the diffusion model above, we emphasized that drift rates, boundaries, and nondecision time are independent components of processing and that this explains why an individual's accuracy cannot be predicted from his or her speed and vice versa. If the components are statistically independent, then it follows that correlations among these parameters should not be large. Ratcliff, Thapar, and McKoon (2010) showed this for numerosity discrimination, item recognition, and lexical decision. For the three parameters for the three tasks, the nine correlations ranged from -0.26 to 0.05. For the three experiments reported here, averaging across the three groups and all the conditions in the experiments, boundary separation and nondecision time correlated 0.28, boundary separation and drift rate correlated -0.33, and nondecision time and drift rate correlated -0.18. Two of these correlations are a little higher than those reported in Ratcliff et al., but are still not large. They show that the model parameters are relatively independent and they are capturing different aspects of the data.



Fig. 3. Plots of accuracy and the .1, .5 (median), and .9 response time (RT) quantiles for data (x-axis) and predicted values from fits of the diffusion model (y-axis) for correct responses for college students. The o's are for data with 25 or more observations and the x's are for between 7 and 24 observations.

Discussion

Summary of results

The data show similar levels of accuracy for older adults and college students but ABLE students were less accurate. The older adults were slower than the college students and the ABLE students were slower than the older adults. For the difference between the matching and mismatching sentences, accuracy and median RTs were similar for the three groups, but these cannot be used to interpret how well the three groups encoded contextually relevant meanings because of the scaling problem mentioned above: large differences in their baseline RTs. This is one of the reasons a model is needed to integrate accuracy and RTs into one, underlying, measure of differences among experimental conditions.

The diffusion model accounted for these effects. Boundary separation and nondecision time increased from college students to older adults to ABLE students. Overall, drift rates were similar for the older adults and college students, but lower for the ABLE students. The difference in drift rates between matching and mismatching sentences was similar for the college students and older adults, but smaller for the ABLE students. The latter might be because the ABLE students were less likely to encode contextually relevant meanings or because the baseline drift rates were different for them. However, the main point here is that the ABLE students did show encoding of contextually relevant meanings.

Why use the diffusion model?

The power of the model has been illustrated by its applications to studies with older adults (ages 65-90; Ratcliff, Thapar, & McKoon, 2001, 2003; Ratcliff et al., 2010, 2011; Ratcliff, Thompson, & McKoon, 2015; Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff et al., 2004; Thapar, Ratcliff, & McKoon, 2003). We discuss Ratcliff et al.'s (2010) study as an example. There were three tasks, numerosity discrimination, item recognition memory, and lexical decision, and three groups of participants, college students, 60-74 year olds, and 75-90 year olds. There were no differences among the three groups in accuracy for numerosity discrimination and item recognition; this would lead to the conclusion that there was no age deficit in these tasks. For lexical decision, the college students' accuracy was lower than that of the older adults; this would lead to the conclusion that they showed a deficit relative to the older adults. In contrast, RTs increased from the college students to the two groups of older adults for all three tasks, which would lead to the conclusion that there was an age deficit.

As we pointed out above, the diffusion model provided a resolution of this contradiction and in so doing, contradicts the long-held view that older adults have deficits in the information upon which decisions are made (for these tasks). Their drift rates were not significantly different from those of the college-age adults (except for a lower value for the college students relative to the older adults in lexical decision, see the same result in Ratcliff et al. (2004)). Drift rates are largely related to accuracy, and boundaries and to a lesser extent nondecision times are



Fig. 4. Plots of accuracy and the .1, .5 (median), and .9 response time (RT) quantiles for data (x-axis) and predicted values from fits of the diffusion model (y-axis) for correct responses for older adults The o's are for data with 25 or more observations and the x's are for between 7 and 24 observations.

largely related to RTs. Accordingly, RTs increased across the groups as the separation of the boundaries increased and nondecision times became longer. The longer nondecision times for the older adults may reflect a deficit in stimulus encoding, response output, memory access, and/or translation of the stimulus representation to a decision-relevant representation. In contrast, boundary separation is under the control of the participant. It is set according to the preference of the participant and older adults set it more conservatively, reflecting their intent to make as few errors as possible (Starns & Ratcliff, 2010).

The model has also accounted for scaling effects across the differences between the conditions of experiments. One way to show this is by simulation: McKoon and Ratcliff (2012) used simulations of priming effects in lexical decision. The primed/unprimed difference in drift rates was held constant at .01 (0.3 and 0.2, respectively) and boundary separation was either 0.08 (about the smallest usually observed) or 0.25 (about the largest usually observed). With the smaller separation, the difference between primed and unprimed mean RTs was 9 ms. With the larger boundary separation, it was 77 ms. In the simulations, the accuracy difference between the conditions was about 9% for young adults and about 10% for older adults.

McKoon and Ratcliff (2012) also showed scaling effects with empirical data. College-age and 60–90 year olds were given lists of pairs of words to be studied, with the two words in a pair being related or not related to each other. After each list, the individual words of a pair were tested, one immediately following the other, for item recognition ("was this word in the study list?"). If the two words of a pair had been related, responses were faster than if they had not been related, and this difference was about the same for the older participants as the collegeage ones (though accuracy differences were larger for college age adults than older adults – about 6% versus 3%). The RT results suggest that memory for the relationship between the words of pairs was as good for the older adults as the college-age ones. However, application of the diffusion model showed that this was not the case. Instead, drift rates decreased with age. The finding that the RT difference between words from related pairs and words from unrelated pairs was not affected by age came about because the older adults set their separation between boundaries larger than did the college-age adults. (McKoon and Ratcliff showed a number of other results that disentangle accuracy and RT differences as a function of age through diffusion model analyses.)

The diffusion model and Experiments 1, 2, and 3

Experiment 1 and McKoon and Ratcliff's 1988 experiment both show college students encoding meaning as it is contextually relevant, with different matching/mismatching stories in this experiment than McKoon and Ratcliff's and more of them (80 vs 40). The diffusion model was not applied to the data from the earlier experiment and so a matching-mismatching effect on drift rate was not separated from participants' boundary settings or nondecision times. In the experiment here, application of the model showed a significant effect on drift rates: "tomatoes are red" was better encoded in memory when its story discussed painting tomatoes than when it discussed rolling them. The model was tested against the data for the older adults and the ABLE students in addition to the data for the college students. In all three cases, the model fit the data well and did so under all the usual constraints (e.g., right-skewed RT distributions; effects of difficulty explained by drift rates and not boundaries or nondecision times; relative speeds of correct and error responses).

As with the aging studies discussed above, the model found contextually relevant encodings by separating the speed/accuracy settings that an individual adopts from the information that he or she encodes. For development of future theories about reading comprehension, it is essential to know to what extent individuals encode textual information- this is not possible without separating away individuals' choices of speed/accuracy settings.

The model also resolved scaling issues. The median RTs for the mismatching test sentences were 1167 ms, 1482 ms, and 1686 ms, and the differences between matching and mismatching test sentences were 79, 86, and 111 ms, for the college students, older adults, and ABLE students, respectively. From these differences, it would appear that the ABLE students showed a somewhat larger matching-mismatching effect than the other two groups (although this difference was not significant). However, the RTs are quite different for the three groups. Transforming them (and accuracy) into drift rates, the pattern was reversed; whereas the matching-mismatching difference in RTs was larger for the ABLE students than the other two groups, the difference in drift rates was smaller, 0.074 for the ABLE students and 0.123 and 0.169 for the college students and older adults, respectively.

Two factors explain this. First, the ABLE students set their boundaries farther apart than the college students and the older adults. Fig. 1C shows how moving boundaries farther apart increases the difference in RTs between two conditions. Second, baseline drift rates decreased from the college students and older adults to the ABLE students. Fig. 1D shows how the same difference in drift rates between two conditions gives a larger difference in RTs for lower baseline drift rates than larger ones.

Importantly, the model fit the data separately for each individual participant, which means that, in future research, values of drift rates, boundary settings, and nondecision times can be correlated with scores on tests of other abilities that might affect comprehension such as IQ, age, and level of education, and scores on more comprehensive tests of comprehension such as the TABE (McKoon & Ratcliff, 2016). Discovering the relations among these measures may inform theories of reading comprehension and it may lead to better tailoring of written information to an individual's skills or better tailoring of methods of teaching comprehension skills.

For the model, there were three key results never before demonstrated. One is that it was successful with a sentence verification task; another is that it was successfully used with adults who do not have fluent reading skills; and the third is that it was successful with RTs considerably longer than has been the case for other paradigms to which it has been applied (means as high as 2.5 s in the experiments here as opposed to the typical 500 ms to 1 s). All three of these findings point the way to future uses of the model by many researchers, in particular paradigms in which the test items are more than a single word– pairs of words, phrases, or sentences.

Implications for theories about older adults' comprehension skills

If understanding which aspects of a word's meaning are relevant in a given circumstance is central in reading comprehension (e.g., Perfetti & Stafura, 2015; Cromley & Azevedo, 2007), then it is critical to test it for older adults. The question is whether the ability to encode contextually relevant meanings declines with normal aging, something not investigated previously. The finding of a significant decline would impact theories of reading comprehension and it would have practical implications for how written information should be expressed for older adults. On the other hand, if there were no significant decline, as there

was not in the experiment reported here, then the skills of normal older adults can serve as a benchmark against which the skills of MCI and AD patients can be compared.

Previous studies of context effects with older adults have produced mixed results, some indicating that they do encode meanings in a contextually relevant fashion and some that they do not, some using only RTs as the dependent variable and others using only accuracy as the dependent variable (see citations above). On the basis of this collection of results, it cannot be decided if theories of comprehension for older adults should or should not predict contextually relevant encoding of meanings.

Our study shows that older adults do encode contextually relevant meanings, or at least they do so in the circumstances of the stories we used. Their longer RTs relative to the young adults were due, in the main, to differences in their speed-accuracy criteria settings. Older adults are much more concerned about accuracy than young adults and so set their boundaries farther apart (Starns & Ratcliff, 2010; Starns & Ratcliff, 2012). It should be mentioned here that speed/accuracy settings might be reversed; instead of responding slowly to maximize accuracy, older adults might become frustrated by a task or they might become concerned that they are responding too slowly and so speed up, sacrificing accuracy.

Our finding conflicts with a number of hypotheses, mentioned above, that have been made about older adults' general cognitive deficits. A deficit in short-term memory might mean that a word and the context that determines its meaning are not in short-term memory at the same time and so which meaning of a word is appropriate is not understood. An inability to bind one piece of information to another might have the same effect. A decrease in processing speed might mean that the time spent on a word and its context is not sufficient for them to be jointly encoded. A decrease in the ability to discard irrelevant information might preserve contextually irrelevant meanings. More specifically, our results could conflict with Craik and Byrd's (1982) proposal that older adults are limited in the richness with which they encode words' meanings and therefore limited in the extent to which context determines what they encode. Whether these hypotheses could be specified in enough detail to explain why older adults can encode contextually relevant meaning and whether there are sorts of meanings they cannot encode are open questions.

Implications for theories about low-literacy adults' comprehension skills

In the Introduction, we emphasized that the number of people in the United States who have difficulty comprehending even the most basic textual information is unexpectedly large. Previous research (described above) has been interpreted as showing that these individuals have many sorts of difficulties, including integrating pieces of information with each other and with general knowledge, forming causal connections, and establishing connections between anaphors and their intended referents. Other studies (described above) have addressed the encoding of contextually relevant meaning specifically and they also have been interpreted as showing difficulties for low-literacy readers.

In contrast, our results demonstrate something that low-literacy adults can do: encode contextually relevant meanings. As with the older adults, we believe that the differences between our finding and earlier ones rest on our use of the diffusion model to explain accuracy and RTs jointly and in so doing, separate the elements of decision-making from each other. The participants in our study were considered low-literacy because they were enrolled in ABLE classes. How to define this population more generally is a difficult question which future research should address.

Our findings also again illustrate how accuracy and RTs lead to inconsistent interpretations of data. For both accuracy and median RTs, the matching-mismatching difference in our experiment was about the same for the three groups of participants while their baseline RTs showed large differences; median RTs for mismatching test sentences were 1482, 1732, and 1167 ms for the older adults, ABLE students, and college students, respectively.

Another way to describe our result is that the ABLE students knew more than their accuracy scores indicated– their matching-mismatching effect in accuracy was only 2%. This is especially salient because it is accuracy that has been used to observe correlations between comprehension skills and such individual difference measures as IQ, age, and scores on global tests like the TABE.

Summary

We reiterate here how important it is to understand what cannot be concluded from accuracy data alone or from RT data alone. Individuals who perform with the same speed may have differences in accuracy, and therefore differences in the information underlying their performance, and individuals with the same accuracy may have differences in speed, and therefore differences in the information underlying their performance. Measuring accuracy alone or RTs alone will almost certainly result in misleading interpretations of data (e.g., Ratcliff et al., 2010, 2011; Ratcliff et al., 2015).

We hope that the outlook on comprehension that we have adopted for this study, the focus on a particular kind of information necessary for full comprehension and a model-based approach to interpreting data, will be useful in identifying comprehension difficulties in ways that can better evaluate older adults' reading skills and lead to better teaching methods for low-literacy adults.

Acknowledgments

This work was supported by the Institute for Educational Sciences (grant number R305A120189) and the National Institute on Aging – USA (grant R01-AG041176).

References

- Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: Knowledge, abilities, and nonability traits. Psychology and Aging, 14, 314–330.
- Allen, P. A., Sliwinski, M., & Bowie, T. (2002). Differential age effects in semantic and episodic memory, Part II: Slope and intercept analyses. *Experimental Aging Research*, 28, 111–142.
- Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles- A problem of polysemy. *Cognitive Psychology*, 7, 167–180.
- Anderson, R. C., Pichert, J., Goetz, E., Schallert, D., Stevens, K., & Trollip, S. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior*, 15, 667–679.
- Arbuckle, T. Y., Vanderleck, V. F., Harsany, M., & Lapidus, S. (1990). Adult age differences in memory in relation to availability and accessibility of knowledge-based schemas. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 305–315.
- Badham, S. P., Estes, Z., & Maylor, E. A. (2012). Integrative and semantic relations equally alleviate age-related associative memory deficits. *Psychology and Aging*, 27, 141–152.
- Baer, J., Kutner, M., & Sabatini, J. (2009). Basic reading skills and the literacy of america's least literate adults: Results from the 2003 national assessment of adult literacy (NAAL) supplemental studies (NCES 2009-81). Washington, D.C.: National Center for

Education Statistics, Institute of Education Sciences, U.S. Department of Education. Balota, D. A., & Duchek, J. M. (1988). Age-related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging*, *3*, 84–93.

- Barnes, M. A., Ahmed, Y., Barth, A., & Francis, D. J. (2015). The relation of knowledgetext integration processes and reading comprehension in 7th- to 12th-grade student. *Scientific Studies of Reading*, 19(4), 253–272.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, 10, 82–93.
- Barton, S. B., & Sanford, A. J. (1993). A case-study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition*, 21, 477–487.
- Bowles, R. P., Grimm, K. J., & McArdle, J. J. (2005). A structural factor analysis of vocabulary knowledge and relations to age. *Journals of Gerontology: Series B. Psychological Sciences and Social Sciences*, 60, P234–P241.
- Bowles, R. P., & Salthouse, T. A. (2008). Vocabulary test format and differential relations to age. Psychology and Aging, 23, 366–376.
- Bowyer-Crane, C., & Snowling, M. J. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology*, 75, 189–201.
- Brod, G., Werkle-Bergner, M., & Shing, Y. (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral*

Neuroscience, 7, 139.

Burke, D. M., & Harrold, R. M. (1988). Automatic and effortful semantic processes in old age: Experimental and naturalistic approaches. In L. L. Light, & D. M. Burke (Eds.). Language, memory, and aging (pp. 100–116). New York: Cambridge University Press.

- Burke, D. M., & Peters, L. (1986). Word associations in old age: Evidence for consistency in semantic encoding during adulthood. *Psychology and Aging*, 1, 283–292.
- Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik, & T. A. Salthouse (Eds.). *The handbook of aging and cognition* (pp. 373–443). (3rd Edition). New York: Psychology Press.
- Burke, D. M., White, H., & Diaz, D. L. (1987). Semantic priming in young and older adults: Evidence for age constancy in automatic and attentional processes. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 79–88.
- Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, 11, 489–503.
- Cain, K., & Oakhill, J. V. (2006). Profiles of children with specific reading comprehension difficulties. British Journal of Educational Psychology, 76, 683–696.
- Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29*, 850–859.
- Cain, K., Oakhill, J. V., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, 96, 671–681.
- Cattell, R. (1963). Theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 54, 1–22.
- Chiarello, C., Church, K. L., & Hoyer, W. J. (1985). Automatic and controlled semantic priming: Accuracy, response bias, and aging. *Journal of Gerontology*, 40, 593–600.
- Cohen, G., & Faulkner, D. (1983). Word recognition: Age differences in contextual facilitation effects. British Journal of Psychology, 74, 239–251.
- Cornelius, S. W., & Caspi, A. (1987). Everyday problem solving in adulthood and old age. Psychology and Aging, 2, 144–153.
- Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. Philosophical Transactions of the Royal Society of London B, 302, 341–359.
- Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik, & S. E. Trehub (Eds.). Aging and cognitive processes (pp. 191–211). New York: Plenum.
- Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99, 311–325.
- Dagerman, K. S., Macdonald, M. C., & Harm, M. W. (2006). Aging and the use of context in ambiguity resolution: Complex changes from simple slowing. *Cognitive Science*, 30, 311–345.
- Dixon, R. A. (2003). Themes in the aging of intelligence: Robust decline with intriguing possibilities. In R. Sternberg, J. Lautrey, & T. Lubart (Eds.). *Models of intelligence: International perspectives* (pp. 151–167). Washington, DC: American Psychological Association.
- Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E. J. (2011). Diffusion versus linear ballistic accumulation: Different models for response time, same conclusions about psychological mechanisms? *Psychonomic Bulletin & Review*, 55, 140–151.
- Duchek, J. M., & Balota, D. A. (1993). Sparing activation processes in older adults. In J. Cerella, J. Rybash, W. Hoyer, & M.L. Commons (Eds.), Adult information processing: Limits on loss (pp. 383–406). Academic Press.
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44, 491–505.
- Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, 42, 133–141.
- Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, 115, 149–161.
- Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39, 133–146.
- Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, 67, 641–666.
- Garnham, A., Oakhill, J., & Johnson-Laird, P. (1982). Referential continuity and the coherence of discourse. *Cognition*, 11(1), 29–46.
- Gernsbacher, M., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 430–445.
- Goral, M., Spiro, A., Albert, M. L., Obler, L. K., & Tabor Connor, L. (2007). Change in lexical retrieval skills in adulthood. *The Mental Lexicon*, 2, 215–240.
- Greenberg, D. (2008). The challenges facing adult literacy programs. Community Literacy Journal, 3, 39–54.
- Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. Psychology and Aging, 7, 56–64.
- Hannon, B., & Daneman, M. (2004). Shallow semantic processing of text: An individualdifferences account. Discourse Processes, 37, 187–204.
- Henderson, L., Snowling, M., & Clarke, P. (2013). Accessing, integrating, and inhibiting word meaning in poor comprehenders. *Scientific Studies of Reading*, 17, 177–198.
- Hess, T. M. (1985). Aging and context influences on recognition memory for typical and atypical script actions. *Developmental Psychology*, 21, 1139–1151.
- Hopkins, K. A., Kellas, G., & Paul, S. T. (1995). Scope of word meaning activation during sentence processing by young and older adults. *Experimental Aging Research*, 21, 123–142.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement

invariance in aging research. Experimental Aging Research, 18, 117-144.

Howard, D. V., McAndrews, M. P., & Lasaga, M. I. (1981). Semantic priming of lexical decisions in young and old adults. Journal of Gerontology, 36, 707-714.

International Adult Literacy Survey Institute. (2011). Retrieved from the website: < http://www.statcan.gc.ca/pub/89-588-x/4152887-eng.htm > (March 3, 2013).

- Kave, G., & Halamish, V. (2015). Doubly blessed: Older adults know more vocabulary and know better what they know. Psychology and Aging, 30, 68-73.
- Kave, G., & Yafe, R. (2014). Performance of younger and older adults on tests of word knowledge and word retrieval: Independence or interdependence of skills? American Journal of Speech-Language Pathology, 23, 36-45.
- Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. Psychology and Aging, 16, 312-322.
- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. European Journal of Cognitive Psychology, 16, 262-284.
- Kutner, M., Greenberg, E., & Baer, J. (2006). National Assessment of Adult Literacy (NAAL): A first look at the literacy of America's adults in the 21st century (Report No. NCES 2006-470). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Labouvie-Vief, G. (1977). Adult cognitive development: In search of alternative interpretations. Merrill-Palmer Quarterly, 23, 227-263.
- Laing, S., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. Language, Speech, and Hearing Services in Schools, 34, 44-56.
- Laming, D. R. J. (1968). Information theory of choice reaction time. New York: Wiley.
- Laver, G. D. (2009). Adult aging effects on semantic and episodic priming in word recognition. Psychology and Aging, 24, 28-39.
- Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. Psychology and Aging, 8, 34-43.
- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). Neuropsychological assessment (4th ed.). New York: Oxford University Press.

Light, L. L., & Albertson, S. A. (1988). Comprehension of pragmatic implications in young and older adults. In L. L. Light, & D. M. Burke (Eds.). Language, memory, and aging (pp. 133-153). Cambridge: Cambridge University Press.

- Light, L. L., Valencia-Laver, D., & Zavis, D. (1991). Instantiation of general terms in young and older adults.
- Long, D. L., & Golding, J. M. (1993). Superordinate goal inferences: Are they automatically generated during comprehension? Discourse Processes, 16, 55-73.
- Long, D. L., Oppy, B. J., & Seely, M. R. (1994). Individual differences in the time course of inferential processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 1456–1470.
- Lovelace, E. A., & Cooley, S. (1982). Free associations of older adults to single words and conceptually related word triads. Journal of Gerontology, 37, 432-437.
- Madden, D. J. (1988). Adult age differences in the effects of sentence context and stimulus degradation during visual word recognition. Psychology and Aging, 3, 167-172. Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud proce-
- dure. Cognition and Instruction, 21, 251-283. McKoon, G., & Ratcliff, R. (1988). Contextually relevant aspects of meaning. Journal of
- Experimental Psychology: Learning, Memory, and Cognition, 14, 331-343. McKoon, G., & Ratcliff, R. (2012). Aging and IO effects on associative recognition and
- priming in item recognition. Journal of Memory and Language, 66, 416-437.
- McKoon, G., & Ratcliff, R. (2016). Adults with poor reading skills: How lexical knowledge interacts with scores on standardized reading comprehension tests. Cognition, 146, 453-469.
- McKoon, G., Ratcliff, R., & Seifert, C. (1989). Making the connection: Generalized knowledge structures in story understanding. Journal of Memory and Language, 28, 711-734.
- Merrill, E. C., Sperber, R. D., & McCauley, C. (1981). Differences in semantic encoding as a function of reading comprehension skill. Memory & Cognition, 9, 618-624.
- Miller, B., McCardle, P., & Hernandez, R. (2010). Advances and remaining challenges in adult literacy research. Journal of Learning Disabilities, 43, 101-107.
- Miller, L. M. S., Stine-Morrow, E. A., Kirkorian, H. L., & Conroy, M. L. (2004). Adult age differences in knowledge-driven reading. Journal of Educational Psychology, 96, 811-821.
- Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. Psychology and Aging, 7, 257-270.
- Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 520-537.
- Naveh-Benjamin, M., Guez, J., Kilb, A., & Reedy, S. (2004). The associative deficit of older adults: Further support using face-name associations. Psychology and Aging, 19, 541-546.
- Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 826-837.
- Oakhill, J. V. (1982). Constructive processes in skilled and less skilled comprehenders' memory for sentences. British Journal of Psychology, 73, 13-20.
- Oakhill, J. V. (1983). Instantiation in skilled and less skilled comprehenders. The Quarterly Journal of Experimental Psychology, 35, 441-450.
- Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. British Journal of Educational Psychology, 54, 31-39.
- Oakhill, J. V. (1993). Children's difficulties in reading comprehension. Educational Psychology Review, 3, 223-237.
- Oakhill, J. V. (1994). Individual differences in children's text comprehension. In M. A. Gernsbacher (Ed.). Handbook of psycholinguistics (pp. 821-848). New York: Academic

Press

- Oakhill, J. V., & Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. Language and Speech, 29, 25-37.
- Oakhill, J. V., Yuill, N., & Donaldson, M. L. (1990). Understanding of causal expressions in skilled and less skilled text comprehenders. British Journal of Developmental Psychology, 8, 401-410.
- Oakhill, J. V., Yuill, N., & Parkin, A. (1988). Memory and inference in skilled and lessskilled comprehenders. In M. Gruneberg, P. Morris, & R. Sykes (Vol. Eds.), Practical aspects of memory: Vol. 2. Chichester: Wiley.
- OECD (2013). OECD skills outlook 2013: First results from the survey of adult skills. OECD Publishing.
- Park, D. C., & Schwarz, N. (Eds.). (2000). Cognitive aging: A primer. Philadelphia, PA: Psychology Press
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. Scientific Studies of Reading, 11, 357-383.
- Perfetti, C., & Stafura, J. (2015). Comprehending implicit meanings in text without making inferences. In E. O'Brien, A. Cook, & R. Lorch, R. (Eds.), Inferences during reading (pp. 1-18). Cambridge University Press.
- Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.). Precursors of functional literacy (pp. 189-213). Philadelphia: John Benjamins Publishing Company.
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. Scientific Studies of Reading, 18, 22-37.
- Perfetti, C., Yang, C.-L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. Applied Cognitive Psychology, 22, 303-318.
- Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. The Journal of the Acoustical Society of America, 97, 593–608.
- Radvansky, G. A., & Dijkstra, G. (2007). Aging and situation model processing. Psychonomic Bulletin and Review, 14, 1027-1042.
- Radvansky, G. A., Gibson, B. S., & McNerney, M. W. (2014). Working memory, situation models, and synesthesia. American Journal of Psychology, 127, 325-342.
- Ratcliff, R. (1978). A theory of memory retrieval. Psychological Review, 85, 59-108. Ratcliff, R. (2002). A diffusion model account of reaction time and accuracy in a two choice brightness discrimination task: Fitting real data and failing to fit fake but plausible data. Psychonomic Bulletin and Review, 9, 278-291.
- Ratcliff, R. (2008). Modeling aging effects on two-choice tasks: Response signal and response time data. Psychology and Aging, 23, 900-916.
- Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. Journal of Experimental Psychology: Human Perception and Performance, 40, 870–888.
- Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the twochoice diffusion model Decision 2 237-279
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for twochoice decision tasks. Neural Computation. 20, 873-922.
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. Psychology and Aging, 19, 278-289.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. Psychology and Aging, 16, 323-341.
- Ratcliff, R., Thapar, A., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. Perception and Psychophysics, 65, 523-535.
- Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. Journal of Memory and Language, 50, 408-424.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. Cognitive Psychology, 60, 127-157.
- Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. Journal of Experimental Psychology: General, 140, 46-487.
- Ratcliff, R., Thapar, A., Smith, P. L., & McKoon, G. (2005). Aging and response times: A comparison of sequential sampling models. In J. Duncan, P. McLeod, & L. Phillips (Eds.). Speed, control, and age. Oxford, England: Oxford University Press. Ratcliff, R., Thompson, C. A., & McKoon, G. (2015). Modeling individual differences in
- response time and accuracy in numeracy. Cognition, 137, 115-136.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating the parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. Psychonomic Bulletin and Review, 9, 438-481.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. Psychological Review, 106, 261-300.
- Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. Psychology and Aging, 21, 448-465.
- Ronnlund, M., Nyberg, L., Backman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. Psychology and Aging, 20, 3-18.
- Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. Cognitive Psychology, 15, 346-378.
- Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. British Journal of Psychology, 84, 171-199.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. Psychological Review, 103, 403-428.
- Salthouse, T. A. (2009). When does age-related cognitive decline begin? Neurobiology of Aging, 30, 507-514.
- Salthouse, T. A. (2014). Quantity and structure of word knowledge across adulthood. Intelligence, 46, 122-130.
- Schaie, K. W. (1970). A reinterpretation of age-related changes in cognitive structure and functioning. In L. Goule, & P. Baltes (Eds.). Life span developmental psychology: Research and theory (pp. 485-507). New York: Academic Press.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding. Hillsdale, NJ: Erlbaum.

- Schretlen, D., Pearlson, G., Anthony, J. C., Aylward, E. H., Augustine, A. M., Davis, A., et al. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society*, 6, 52–61.
- Scialfa, C. T., & Margolis, R. B. (1986). Age differences in the commonality of free associations. *Experimental Aging Research*, 12, 95–98.
- Seifert, C. M., McKoon, G., Abelson, R. P., & Ratcliff, R. (1986). Memory connections between thematically similar episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12, 220–231.*
- Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, 20, 539–548.
- Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, 5, 377–390.
- Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin and Review*, 19, 139–145.
- Staudinger, U. M., Cornelius, S. W., & Baltes, P. B. (1989). The aging of intelligence: Potential and limits. Annals AAPSS, 503, 44–58.
- Stine-Morrow, E. A., Miller, L. S. M., & Hertzog, C. (2006). Aging and self-regulated language processing. *Psychological Bulletin*, 132, 582–606.
- Tabossi, P. (1982). Sentential context and the interpretation of unambiguous words. *Quarterly Journal of Experimental Psychology*, 34, 79–90.
- Tabossi, P., & Johnson-Laird, P. N. (1980). Linguistic context and the priming of semantic information. Quarterly Journal of Experimental Psychology, 32, 595–603.
- Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, 18, 415–429.
- Todaro, S., Millis, K., & Dandotkar, S. (2010). The impact of semantic and causal relatedness and reading skill on standards of coherence. *Discourse Processes*, 47, 421–446.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592.

- Uttl, B. (2002). North American adult reading test: Age norms, reliability, and validity. Journal of Clinical and Experimental Neuropsychology, 24, 1123–1137.
- Verhaeghen, P. (2003). Aging and vocabulary scores: a meta-analysis. Psychology and Aging, 18, 332-339.
- White, K. K., & Abrams, L. (2004). Phonological priming of preexisting and new associations in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 645–655.
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, 54, 39–52.
- Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*, 14, 133–145. Wingfield, A., Aberdeen, J. S., & Stine, E. A. (1991). Word onset gating and linguistic
- Wingheid, A., Aberdeen, J. S., & Stine, E. A. (1991). Word onset gating and inguistic context in spoken word recognition by young and elderly adults. *Journal of Gerontology*, 46, P127–129.
- Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech*, 37, 221–235.
- Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you mean! Event-related potentials reveal multiple aspects f context use during construction of message-level meaning. *NeuroImage*, 62, 356–366.
- Wlotko, E. W., Federmeier, K. D., Fabiani, M., & Gratton, G. (2011). It takes two: The event-related optical signal reveals distinct contributions from the cerebral hemispheres to language comprehension. *Psychophysiology*, 48, S99 S99-S99.
- Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Agerelated differences in the use of sentential context. *Psychology and Aging*, 27, 975–988.
- Wlotko, E. W., Lee, C., & Federmeier, K. D. (2010). Language of the aging brain: Eventrelated potential studies of comprehension in older adults. *Language and Linguistics Compass*, 4, 623–638.
- Yang, C.-L., Perfetti, C., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. Written Language & Literacy, 8, 157–181.
- Yuill, N., Oakhill, J. V., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology*, 80, 351–361.